DOCUMENT RESUME

ED 128 433 TM 005 618

AUTHOR Johnson, Richard T.
TITLE Conducting Large-Scale Examinations: A Handbook.
PUB DATE Jul 71
NOTE 165p.

EDRS PRICE MF-$0.83 HC-$8.69 Plus Postage.
DESCRIPTORS Administration; *Admission (School); *Competitive Selection; Costs; *Developing Nations; Examiners; Group Tests; *Guidelines; Higher Education; Item Analysis; Personnel Selection; Pretesting; Publicize; Scoring; Secondary Education; Test Construction; *Testing; *Testing Programs; Vocational Education
IDENTIFIERS Test Security

ABSTRACT

In developing countries, for the most part, traditional methods of selecting students for secondary schools, technical and vocational institutions, teacher training colleges, and universities are unable to ensure that the most capable applicants receive the limited opportunities available. To remedy this situation, the author has assembled and incorporated into this handbook a collection of testing procedures which he has found useful in large-scale testing programs in developing countries. In his discussion, the term "large scale" describes examinations taken by at least a thousand candidates on a single schedule date. Nevertheless, most of the chapters are applicable to examinations given on a smaller scale. The handbook may also be valuable to school headmasters and teachers who are interested in the workings of a testing organization, a body that often exerts great influence on their lives and those of their pupils. The handbook is divided into these sections: (1) Publicity, (2) Registration, (3) Item Preparation, (4) Pretesting, (5) Editing, (6) Reproduction, (7) Distribution/Administration, (8) Scoring/Processing/Reporting, (9) Overall Security, and (10) Costs. (RC)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED128433

Conducting Large—Scale Examinations: A Handbook

Richard T. Johnson

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

TM005 618

2

July 1971

Virginia Polytechnic Institute and State University
11440 Issac Newton Square, North
Reston, Virginia 22090

TABLE OF CONTENTS

0	Preface	
	background	0-1
	definition of "large-scale"	0-2
	assumptions	0-2
	usage of the handbook	0-3
1	Publicity	
	brochures	1-2
	advertisements	1-3
	candidate information booklets	1-3
	meetings, lectures	1-4
	sample candidate booklet	1-5
	sample headmaster booklet	1-11
2	Registration	
	candidate forms	2-2
	receipts/admission forms	2-3
	centers & allocation	2-3
	rosters, packing lists, print orders	2-3
3	Item Preparation	
	specifications	3-2
	levels of thinking	3-9
	training of item writers	3-15
	commissioning items	3-17
	notes on item writing	3-21
4	Pretesting	
	rationale	4-2
	pretryout	4-4
	trial sample	4-6
	directions	4-6
	format	4-6
	timing	4-7
	motivation	4-10
	security	4-10
	administrative arrangements	4-13
	answer sheets	4-13
	item analysis	4-13

5 Editing
rewriting 5-2
item file 5-2
arrangement of items 5-3
editing committees 5-6
6 Reproduction
genefal methods 6-2
tests 6-3
directions, accessory materials 6-6
answer sheets 6-6
security 6-9
7 Distribution/Administration
counting, numbering 7-2
packing 7-2
transport 7-3
administration 7-3
return of materials 7-6
timing 7-7
security 7-8
sample instructions 7-9
8 Scoring/Processing/Reporting
prescanning 8-2
scoring 8-2
editing 8-7
cross-footing, combining 8-8
transformations 8-10
listing 8-12
certificates 8-14
distribution of results 8-16
quality control 8-17
statistics, reports 8-18
retention of data 8-18
9 Overall Security
security areas 9-1
safes 9-6
receipt of materials 9-7

9 (continued)

storage 9-9

packing 9-9

shipping 9-11

test centers 9-13

data processing department 9-15

typing, copying, printing 9-16

confidential files 9-17

security breaches 9-18

10 Costs

survey 10-1

each stage 10-1

totals 10-3

FIGURES

0.1	Overall diagram of the conduct of large-scale examinations	0-4
2.1	Entry form for the WAEC/London GCE	2-4
2.2	(reverse side)	2-5
2.3	Sample packing list	2-6
3.1	Example of a weighted blueprint	3-16
3.2	Instructions to item writers for submitting items	3-18
4.1	Edge stripes to prevent loss of single pages	4-12
4.2	Horizontal and vertical formats for candidate ident.	4-13
4.3	Sample item analysis data	4-19
5.1	Completed item card	5-4
5.2	Plot of item statistics	5-4
6.1	Comparison of four common methods of reproduction	6-3
6.2	Precoded answer sheet	6-8
6.3	Answer card	6-8
7.1	Certificate of examination conduct	7-5
7.2	Typical examination hall layout for forty candidates	7-11
8.1	Overprinting a scoring key	8-4
8.2	A computer-printed frequency distribution	8-11
8.3	Results listing	8-13
8.4	Public results listing	8-15
8.5	A Primary School Leaving Certificate	8-16
9.1	Office organization	9-2
9.2	Security area floor plan	9-2
9.3	Stores register	9-4
9.4	STO inventory control	9-4
9.5	Shipping register	9-4
9.6	Access register	9-4
9.7	Key register	9-4

PREFACE

In Nigeria, a new pottery factory which was about to open advertised 100 positions; over 20,000 people applied. When the Polytechnic College of the University of Malawi was ready to begin instruction, applications for the 60 places were received from 1,350 candidates. In West Africa, only ten to eleven per cent of primary school leavers can hope to begin secondary school, and only one per cent of secondary-school-age children are actually attending school at any level.

The situation is similar in hundreds of countries, and in each, some authority must decide which particular individuals will receive more training, a job, or promotion. Although such selection is a problem anywhere, the sheer magnitude of the numbers involved in developing nations magnifies and compounds it.

Traditional methods of selecting students for secondary schools, technical and vocational institutions, teacher training colleges, and universities clearly offer little help, since they are unable to ensure that the most capable applicants receive the limited opportunities. When these procedures are used in countries where educational facilities are seriously circumscribed, precious resources are squandered.

In many African countries, it has been demonstrated that adding objective tests of achievement and aptitude to existing selection procedures can greatly increase the number of above-average performers within the educational system. Tests of aptitude have also proved useful in selecting applicants for training courses, as well as for jobs in public service and in the private sector.

Any government or institution which attempts to use modern methods of testing, however, is faced with many potential pitfalls, unless it has access to some general guidelines devised by people who have worked in similar situations. General guidelines for testing programs are rare, and usually based on experience in America and Europe, experience which is seldom directly applicable in other cultures. To remedy this situation, I have assembled and incorporated in this handbook a collection of testing procedures which have been found useful in developing countries, together with experiences which, although in retrospect have their humorous aspects, were agonizing at the time. I have done this in the fervent hope that at least some disasters of the past will be averted in the future.

The justification for this handbook is that it deals with problems specific to developing countries, and that it includes material not found in commonly available sources. Most of the material assembled here resulted from direct experience in the Pacific, as well as in West and Southern Africa: Western Samoa, Nigeria, Ghana, Sierra Leone, The Gambia, Malawi, Botwana, Lesotho, and Swaziland. Since 1965 I have worked with the West African Examinations Council and with the Regional Testing Resource and Training Centre in Southern Africa on problems of test development. In deference to the nations involved, I have seldom specified the location of a particular faux pas, although the various ministries were wise enough to know that publication of these procedures and errors might assist others.

In the following discussion, the term "large-scale" describes examinations taken by at least a thousand candidates on a single scheduled date. Nevertheless, most of the chapters are applicable to examinations given on a smaller scale. The handbook may also be valuable to school headmasters and teachers who are interested in the workings of a testing organization, a body that often exerts great influence on their lives and those of their pupils. In order to present a unified discussion, I have assumed that:

1. there exists a central office with the major responsibility for conduct of the examination; in this handbook I have called this office the State Testing Office, abbreviated STO, which abbreviation I have often used as a substitute for the editorial "we."
2. the examination consists mainly of multiple choice questions. For pedagogical purposes I refer to three separate papers--A, B, and C--but also deal with a language essay on occasion. In spite of this focus, the handbook should be useful in those countries where only traditional essay examinations are given, since most of the problems are identical.
3. the major purpose of the examination is to select a proportion (often less than .25) of the possible candidates for further treatment. The treatment could be secondary education, apprenticeship training, working at a factory, promotion in the civil service, etc. Actually, most of the material here is centered upon secondary school selection, but the methods are applicable to other levels and types of assessment.
4. the candidates, supervisors, invigilators (proctors), and users of the test results may speak English as their second language. Consequently,

I have tried to refrain from using the usual testing jargon in all material which such individuals will use. In some instances, such as the Candidate Information Booklet, the English used, although not always the most correct, has been found to give the best results in actual use.

5. the users of this handbook are acquainted with the elementary concepts of psychological measurement as presented in most basic texts. In spite of this, some points which I present may be obvious to the reader. If too many are, I offer my apologies.

Although I appreciate the distinction made in some countries between the use of the term "examination" as applying to measures of achievement, and "test" as applying to published measures of aptitude, I have used these terms interchangeably.

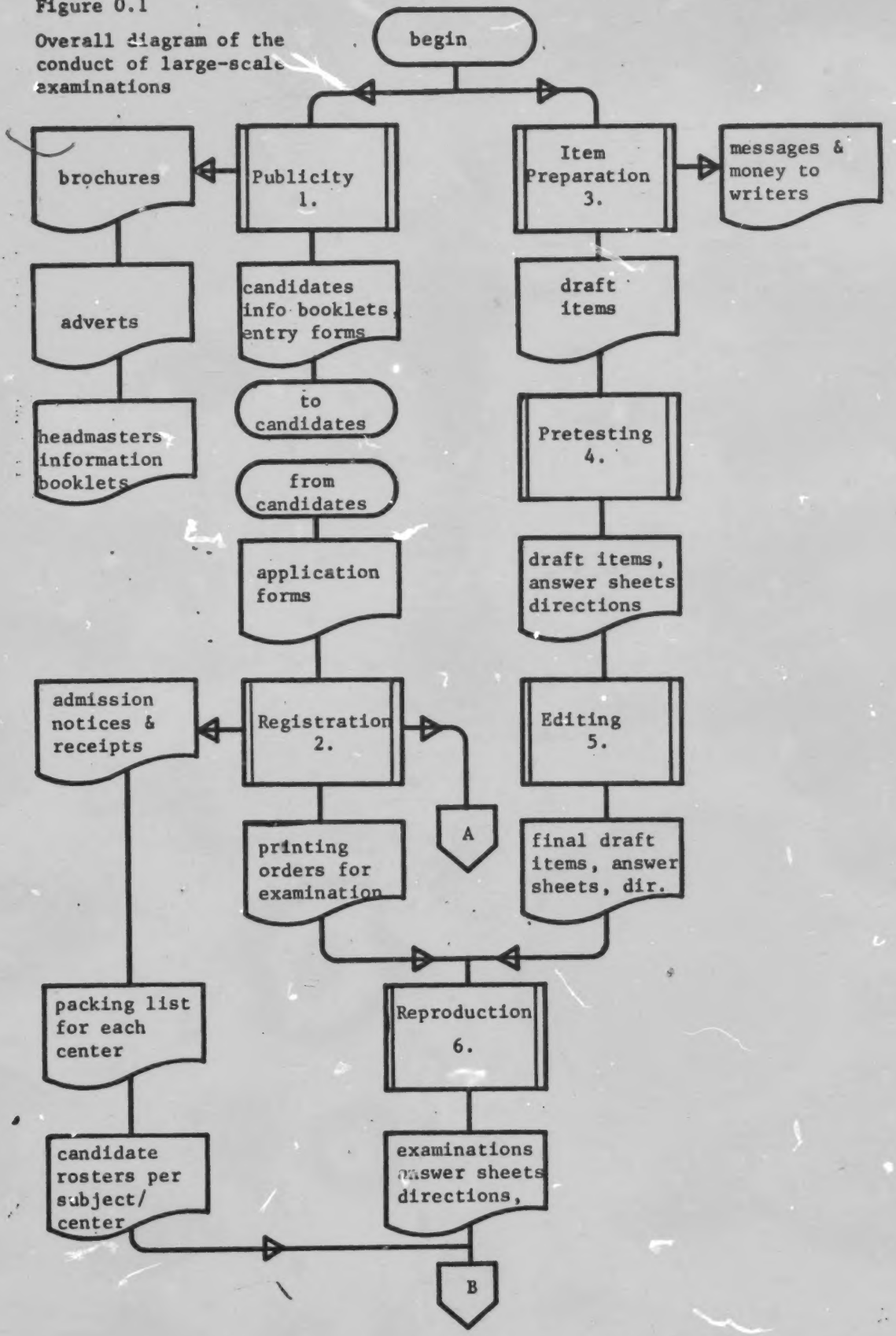
Throughout the handbook, words, phrases, or sentences which are enclosed in brackets [ ] show alternatives which may apply more in one setting than in others, or are infrequently used. Throughout I have tried to give officials general titles for which users could most easily substitute the actual titles of their corresponding local officials.

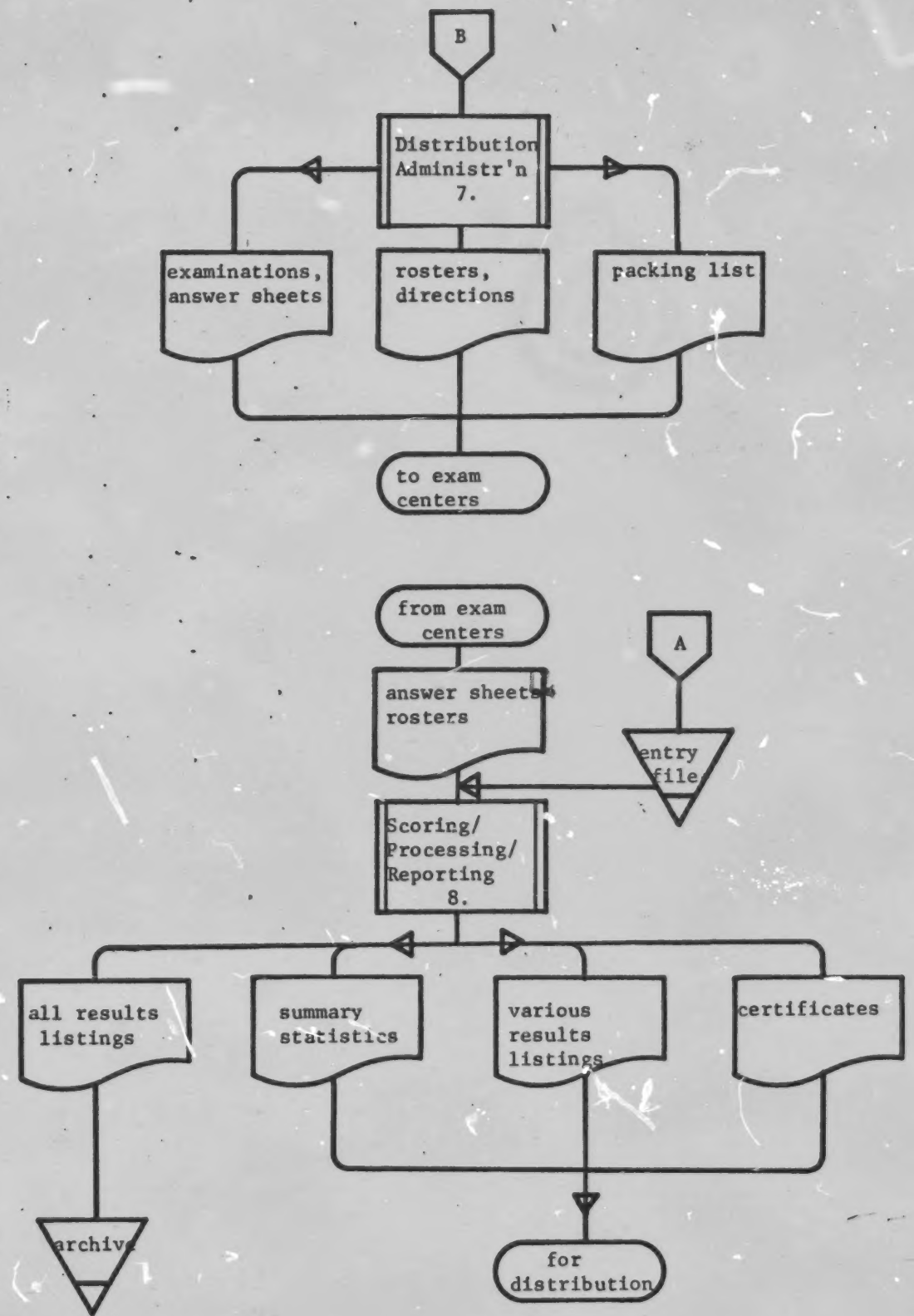
Before he begins reading the actual discussion, the reader may wish to refer to the flowchart of the entire examination cycle on the next two pages in order to get an overall view of its sequence and of the interrelationships among parts. Specific sections of this chart reappear in greater detail at the beginning of those chapters which deal with them.

Grateful appreciation in expressed to Professor Lee J. Cronbach of Stanford University and to Dr. Sonia Johnson for their encouragement and support. Financial assistance for the testing projects in Africa came from the United States Agency for International Development. Neither these persons or institutions, nor any I have left unmentioned, is to be blamed for my mistakes.

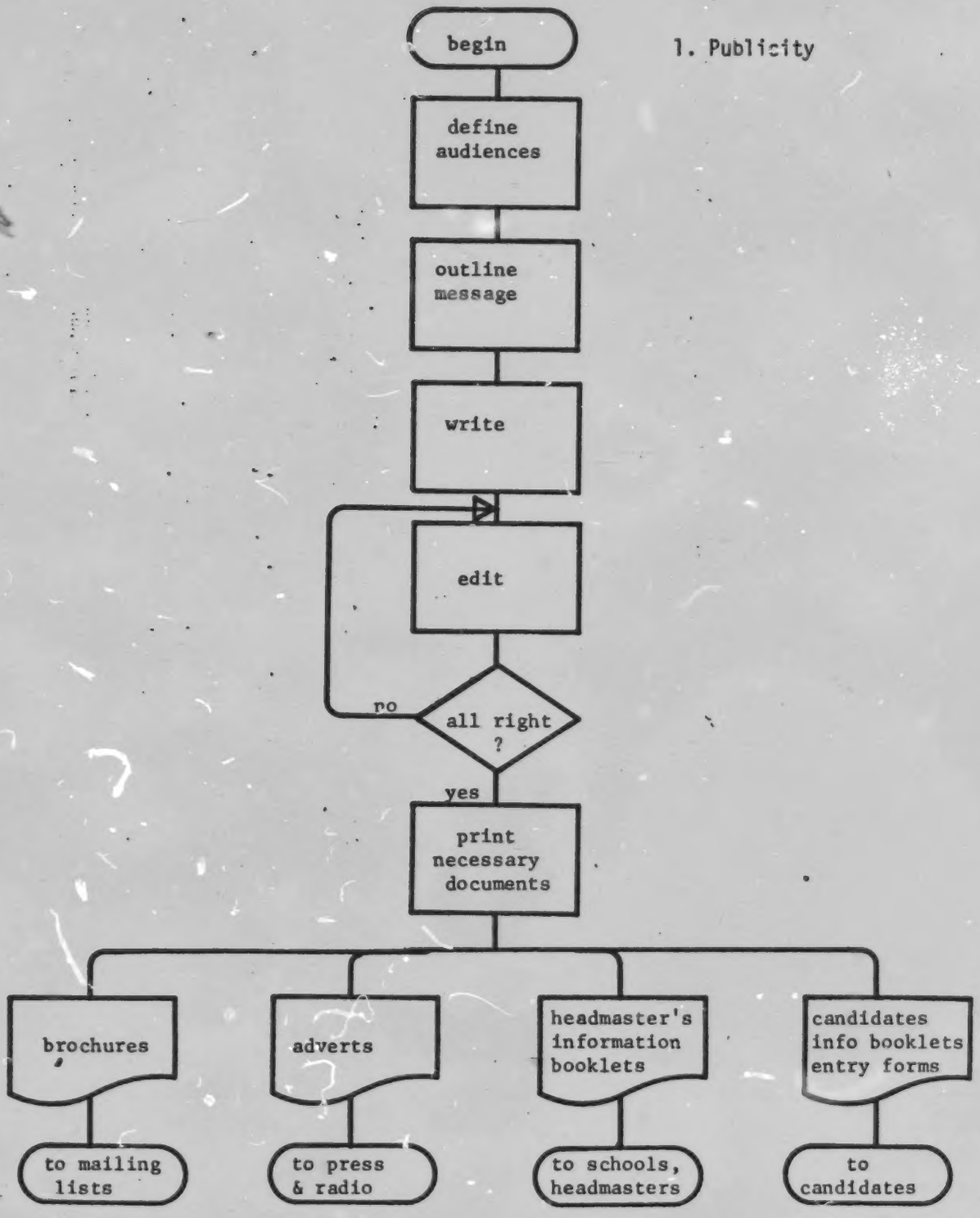
Figure 0.1

Overall diagram of the conduct of large-scale examinations




B
Distribution
Administr'n
7.
examinations,
answer sheets
rosters,
directions
packing list
to exam
centers
from exam
centers
A
answer sheets
rosters
entry
file
Scoring/
Processing/
Reporting
8.
all results
listings
summary
statistics
various
results
listings
certificates
archive
for
distribution

1. Publicity


begin
define audiences
outline message
write
edit
all right ?
no
yes
print necessary documents
brochures
adverts
headmaster's information booklets
candidates info booklets entry forms
to mailing lists
to press & radio
to schools, headmasters
to candidates

PUBLICITY

No matter how much a State Testing Office would like to remain inconspicuous, everyone from the personnel officer and the headmaster to the apprehensive candidate and the man on the street will have distinct (but not necessarily accurate) impressions of the scope and accuracy of their work. The major decision which the testing office must take, therefore, is not whether there should be publicity, but whether they should originate it themselves or allow the informal sources to bear the whole responsibility.

The former course is preferable. If informal sources alone are relied upon, the information which reaches the public is often inaccurate and perhaps even derogatory. Failure to publicize information as basic as examination dates and times is not unusual, and causes a great deal of hardship to both candidates and STO officers. For instance, in one country an examination had been given for many years on the last Thursday of November at 08.30. A change was made to an earlier hour with essentially no publicity given the fact, and the many candidates who showed up at the wrong time were not allowed to take the examination. Their government representatives, to whom they complained, were understandably upset and gave instructions to the STO to set another examination immediately. Obviously, this not only required an enormous amount of work and expense on the part of the testing office, but also contributed to unfairness to candidates since it is virtually impossible to set two different examinations of exactly the same difficulty within a short time.

Staff morale is also affected adversely when the public lacks positive information about the STO. If working at the testing office is regarded as having little status, staff members may come to feel that the sooner they find more prestigous jobs, the better. Every time someone with this attitude leaves, the status declines a little further. A constant effort of an administrator to build up the image of the office, and to have staff members given university appointments for the added prestige will increase morale considerably.

There are a number of ways in which information can be disseminated to the candidates as well as to those having responsibilities for administration of the examination and to those who are most likely to use the results. An obvious method is to distribute a descriptive brochure including information about the testing office, its services, the major types of examinations it may offer, and special consulting assistance it may be prepared to give.

Some offices prefer to issue yearly brochures in which they give the timetable for their major examinations. Others, which prepare less frequently a general pamphlet covering all activities but not giving specific dates or times, rely on other modes of publicity to advertise the examination schedule.

No matter which approach is taken, it is wise to insert advertisements in the local press concerning the examination details. Each notice should include the time and places of the examination, who is eligible to take it and for what purpose, and who would be likely to use the examination results. The advertisement should include, in addition, the method of applying to take the examination and the cost.

Notices in newspapers often lead to unexpected improvements. In one West African country, it had long been a practice for headmasters to charge an exorbitant fee to every candidate for the privilege of receiving an entry blank. Thus a candidate had to pay twice, first for the entry blank to apply to take the examination, and second to take the examination itself. This practice lined the pockets of many a headmaster and impoverished many a parent.

The Headmasters' Association had sharply disapproved of this onerous practice for many years but all its attempts to abolish it had met with failure. Publicity finally accomplished what rules and regulations could not. When an advertisement was placed in all national newspapers announcing that candidates could obtain free entry blanks from their headmasters or from the STO, the practice was almost completely eliminated.

The final blow to this corruption was the distribution of candidate information booklets. On the front page were the words: "distributed free to all candidates" and on the back was an entry blank.

The story has a less happy sequel. When a supply of these booklets was sent by lorry to the adjacent country for distribution, some headmasters sent word to their colleagues in the other country ahead of the lorry, and the booklets were intercepted. After the notice of free distribution was blotted out, the truck was allowed to continue, and so the practice of selling entry blanks even now continues in that country.

Candidate information booklets, nevertheless, remain the best publicity device available. Since the information reaches those who are most vitally affected by the examination, every detail is devoured, so the booklet should contain information which candidates most need; descriptions of the papers in the test and of the types of questions, instructions for marking the answer

sheet and for applying to take the examination, and an explanation of what will be done with the results. A typical information booklet is shown on page 1-5;it provides candidates with opportunities to practice their test-taking skills and to obtain feedback on their performance.

In the example, the last page of the booklet is perforated so that the candidate may remove it and use it as a sample answer sheet. An application to take the examination is printed on the reverse side. After the candidate successfully fills out the sample answer sheet, he completes the entry blank on the opposite side and gives it to his headmaster.

The first time such an approach is used in a country, the testing office staff might well collect a sample of the entry blanks from the schools and inspect them to determine if the candidates are understanding the directions and marking the answers correctly.

The candidates are not the only ones who are vitally interested in the examination; the headmasters also view it as crucial. Consequently distribution of a headmaster's information booklet similar to the sample on page 1-11 is a valuable service the STO can perform.

Another method of providing publicity for the examinations and for the testing office is through lectures and speeches presented by staff, and by their attendance at various meetings related to their work. Sometimes these publicity techniques, although considered inferior to the usual radio and newspaper advertising, pay greater dividends. This is true because the groups which hear the lectures and attend the meetings are already very interested in the subject and spread the information readily.

One other aspect of publicity might be mentioned. No matter how professional the testing office personnel are, or how tried and true the methods of administering and processing the examination might be, mistakes are inevitable. These result in a tremendous amount of adverse publicity, and any testing office would do well to donsider producing a fair amount of favorable publicity in advance to counteract it. The West African Examinations Council has gone so far as to hire an officer who has the sole responsibility of producing regular publicity releases to all the news media. Since the officer consequently understands almost all the important aspects of the testing office, he is in a good position to deal with any complaints which may reach the press in regard to results, general mistakes, security breaches, or poor choice of test content.

Note: This booklet is distributed free to all candidates.

CANDIDATE INFORMATION BOOKLET

This booklet has been written to help you to understand how to answer questions on the examination you will soon be taking. Be sure to read these instructions before the examination day. Practice on this paper and get used to filling in the answers correctly. Doing so will help you do your best on the real examination.

Within a short time, you and thousands of candidates like you will take a special test. Three kinds of papers will be included. One of them is set to test how much you have learned in English during your past years at school; another to measure what you have learned in arithmetic. The third paper is a general paper which should show how easily you can learn English, arithmetic, and other subjects in the future.

Each examination paper will have many different questions and each question will have four or five choices, one of which will be the correct answer. These choices will be marked A B C D or A B C D E. You will be asked to select the choice you think answers the question best. The answers will be marked on a special answer sheet [card]; you will be told how to do this later.

First of all, each of you will be given an examination number which will have to be filled in on your answer sheet. For example, your number could be 906 (of course during the real examination it will be different but use this for the sample). This number is to be marked on every answer sheet which you use. In the back of this booklet is a sample answer sheet to help you learn how to fill in your examination number and how to mark your answers. Please tear off the last page of this booklet. There are spaces for marking answers to test questions on one side, and an entry form to take the examination on the other side. When you have finished writing your examination number on the sample answer sheet on page 1-9 , try the sample test questions in this booklet, and then compare your whole sample answer sheet with the correct one on page 1-9 . As soon as you are sure that you understand the way to mark your examination number, and your answers, fill in the entry form and give it to your Headmaster.

Now, take the sample examination answer sheet and turn it sideways. You will see blank squares along the top under the heading "Student Number." Fill in your sample examination number in the squares like this -- [9][0][6] -- note that each box has a single number in it. Then, under the number 9 you will

see small boxes containing numbers. Go down this column until you find number 9 and then, with a pencil (not a pen or ball-pen) black out this box completely, being careful to stay within the lines. Then look down the second and third columns and do the same. If your number includes a zero as our sample examination number did, remember to black out the 0 like any other number. Once you have finished, turn your sheet upright again.

Now that this has been done you are ready to answer the questions. On pages 1-8 and 1-8 are sample problems called Test 1 and Test 2. The answers are also in this booklet but do not look at them yet. First do the problems on Test 1 and compare your answers and way of marking with the correct ones for Test 1 (use the sample answer sheet on page 1-9 and compare it with the correct answers marked on page 1-9 .)

Make your corrections if any. Then do the problems on Test 2 and compare your answers with the correct ones shown on page 1-9 . In every case, read the question first and then the different answers. When you have selected the best answer for that question, look at the answer sheet where you will see small boxes with letters along the same line as the number of your question. Using a pencil, black in the letter which matches the letter of your choice on the question paper. Black the box in completely, without going outside the edges. Do the same for all questions, making sure that you fill in the correct lines which match your question paper. The questions are like some that you will be asked on the examination. Answering them will help you to learn how to show your answers in the right way. This practice examination does not count for any credit.

Points to remember when taking the test:

1. In the examination room, listen very carefully to all instructions given by the supervisor.
2. Answer as many questions as you can, but do not waste time on a question you think you cannot answer. When you have finished all the ones you can answer, you can then try the more difficult ones again.
3. *YOU ARE NOT ALLOWED TO USE INK OR MAKE ANY MARK YOU CANNOT RUB OUT.* You should therefore bring two or more sharpened pencils and a good eraser.
4. All of your answers must be marked on the answer sheet and *not* on the question paper.
5. If you wish to change your answer in any way, you must first completely erase the mark in the box you have blacked in, before filling in another

one. If you leave more than one box for a question blacked in, the answer will not be marked.

6. Remember to read all directions and questions very carefully as each paper will have its own different instructions.

[Many candidates wonder whether or not to guess the answers to questions about which they are not sure. In this test a percentage of the wrong answers is subtracted from the number of right answers as a correction for guessing. Therefore guessing will not help your score, and it takes time. If, however, you are not sure of the correct answer but have some knowledge of the question and a good idea what might be correct, you should try to answer the question.]

The best way to prepare for the test is to work hard at your studies. Pay attention in every class, and do all your assignments.

While you are doing the test, remember to:

1. read each question carefully,
2. think,
3. choose an answer
4. think again to make sure your answer makes sense,
5. work quickly, neatly, carefully, and
6. write your answer on your answer sheet.

You will receive an admission notice telling you where you are to take the test and giving you your examination number. You *must* bring this with you or you will not be allowed to enter the examination hall. If you lose your admission notice, send a letter or telegram to the State Testing Office at once. Your entry form will be checked and you will be sent a letter which you can use in place of your admission notice.

You should arrive at the examination hall on time since you cannot be allowed to enter after testing has begun. In addition, you will not be permitted to continue any test paper beyond the scheduled time.

No visitors are allowed in the examination hall. Family or friends must remain outside until the test is over and the candidates have been dismissed.

Any person who gives or receives help during the test will be asked to turn in his materials and leave the room. The supervisor will return these materials to the State Testing Office with a note of explanation.

Test marks will be reported directly to the schools you have chosen, and the schools will notify you to come in for interview. Please do not ask the State Testing Office or any headmasters for your results.

TEST 1

Choose the word which has the same meaning, or most nearly the same meaning, as the word at the beginning of the line :

1. SINGLE: (A) one (B) many (C) some (D) all (E) few

2. DOG: (A) afraid (B) song (C) animal (D) large (E) fly

In each of the following problems a word is missing. Choose from the five words the one which most suitably completes the problem :

3. The _____ must bend when the wind blows upon it.

(A) ground (B) house (C) path (D) grass (E) sky

4. Boys will become _____ .

(A) infants (B) little (C) clever (D) stupid (E) men

5. Male and boy. Female and _____ .

(A) man (B) woman (C) girl (D) lady (E) daughter

TEST 2

Solve each problem, using the blank spaces on the page for rough work. Then mark the space on your answer sheet corresponding to the one correct answer :

1. In which of the following ways could 168 pencils be packaged for shipping ?

(A) 11 boxes with 18 pencils in each
(B) 14 boxes with 12 pencils in each
(C) 17 boxes with 14 pencils in each
(D) 24 boxes with 12 pencils in each
(E) 28 boxes with 11 pencils in each

2. Which is the least of the following fractions ?

(A) 1/2 (B) 1/3 (C) 2/7 (D) 3/8 (E) 4/9

3. What is the next number in the series ?

2 3 6 7 10 ?

(A) 11 (B) 12 (C) 13 (D) 14 (E) 15

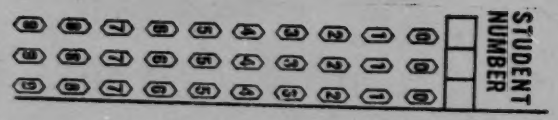
4. A ball team played 37 games and won or tied 24 of them. How many did it lose ?

(A) 3 (B) 12 (C) 13 (D) 61 (E) none of these

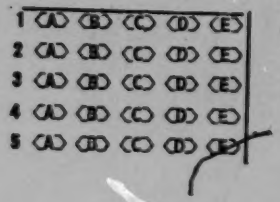
5. How many pounds do 28 feet of uniform wire weigh, if 154 feet weigh 11 pounds ?

(A) 14 (B) 28/11 (C) 11/2 (D) 7 (E) 2

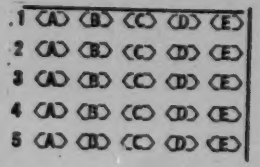
sample answer sheet



Test 1



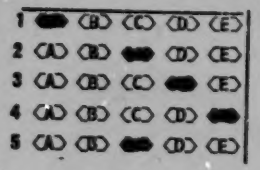
Test 2



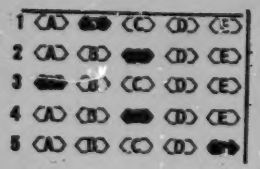
correctly marked answer sheet

Do not look at this one until you have tried the sample test.

Test 1



Test 2



Entry Form

Please print your name as you wish it to appear on the result sheets.

SURNAME (last name) . . . . . .
(in capitals)

FIRST NAMES . . . . . .
(in capitals)

DATE OF BIRTH . . . . . .
(day, month, year)

PLACE OF BIRTH
TOWN OR VILLAGE . . . . . .PROVINCE . . . . . .

SEX Boy ( ) or Girl ( ) (tick one)

SECONDARY SCHOOLS TO WHICH YOU ARE SEEKING ADMISSION

1st choice . . . . . .

2nd choice . . . . . .

3rd choice . . . . . .

SIGNATURE OF CANDIDATE . . . . . .

I agree to the entry of the candidate whose signature is above.

. . . .
(signature or mark of parent or guardian) (witness to signature or mark)

This form should not be sent to the State Testing Office but should be kept at the school.

HEADMASTERS' INFORMATION BOOKLET

Introduction

The purpose of this booklet is to explain to principals, headmasters, teachers, and supervisors what the ABC Examination is like, why and how it is used.

The State Testing Office administered the ABC Examination for the first time in 1960. Successive editions of the test have been given every year since then, and thousands of students are admitted to secondary schools each year on the basis of the tests.

A careful reading of this booklet will give you a clear idea of the test itself, how it was developed, what the scores mean, and how they are used. It should assist you in making sure that every candidate:

1. knows how to fill in his name, student number, and answers on the answer sheet [card] and
2. reads, understands, and works the examples in the Candidate Information Booklet.

The Nature and Purpose of the ABC Examination

The examination is a series of three tests which requires a total of four hours to take. There are appropriate breaks between the tests. These tests are:

Paper 1	Arithmetic	65 minutes
Paper 2	English	70 minutes
Paper 3	General Aptitude	105 minutes

Paper 1 tests knowledge of arithmetical operations, number concepts, and how well the candidate can solve problems he has dealt with in school. Paper 2 measures vocabulary, reading skill, grammar, and ability to understand a given passage. Paper 3, the general aptitude section, measures the candidate's ability to understand and reason with mathematical and verbal symbols and to use them in solving problems. In addition, papers 1 and 2 give an indication of how much the candidate knows already, whereas paper 3 gives an idea of his basic ability to learn.

A separate score is given for each paper because the abilities being measured are different. A candidate may be stronger in one area than in another, and so the separate scores give a clearer picture of his abilities than a single score would.

Development of the ABC Examination

Before the test was published, its designers had to be sure that it was equally fair for candidates from all parts of the nation, and that it could be used to test pupils with large differences in intelligence.

Many years of hard work go into preparing the final edition, and many questions have to be answered: Were the earlier editions accurate? What kinds of new questions should be used? What have the principals, supervisors, headmasters, teachers, and all others who have used the test said about it?

Of course, the content of the new ABC is given the most careful attention. For each paper of the examination, hundreds of questions are written by subject matter and test technique specialists at the State Testing Office. For each of these questions, four or five possible answers are given. The candidate is instructed to choose the one answer he thinks is correct, or at least the best one. The author of the question must word it and the answer choices so that the students who really know the right answer can be separated from those who do not. This is not so easy as it sounds.

After the best questions have been chosen and edited, they are tried out on groups of students as much as possible like those who will be taking the final edition of the examination. After these students have taken the tests, and their answers are studied, more questions are discarded. Usually only ten per cent of the best questions from the first draft are included in the final edition. And very often these questions have been rewritten so many times that they scarcely bear any resemblance to their original versions.

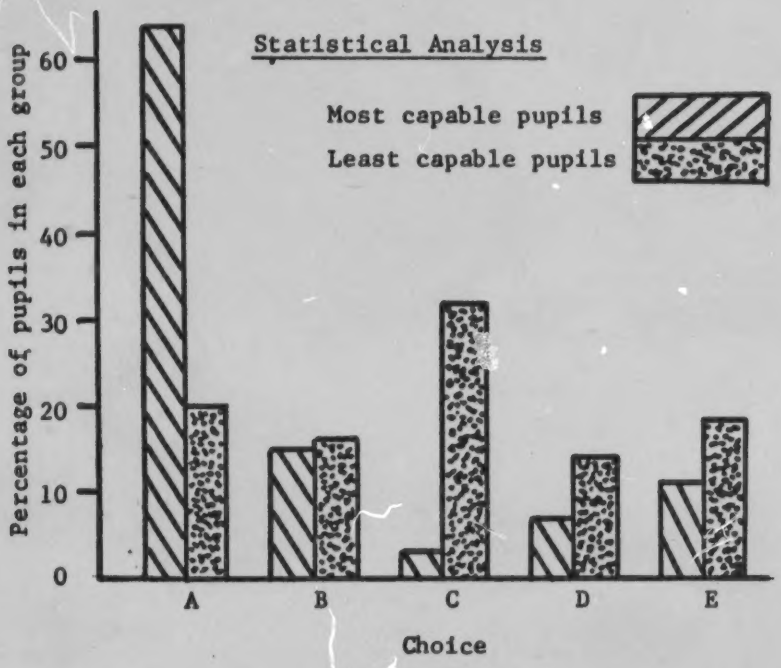
After the trial testing, the best questions are selected by statistical methods. In this process, the *difficulty* of each question is first determined by how many students answer it correctly. The more students who answer a question correctly, the easier the question is judged to be. Similarly, the attractiveness of each answer choice can be determined by the number of pupils who choose it. Any question which is too easy or too difficult is of no value since it cannot separate the able from the less able candidates and therefore cannot fulfill the purpose of the examination.

The *power* of a question to differentiate among pupils of different abilities is a more complex matter, and is usually determined by elaborate statistical methods and a critical evaluation by specialists in the subject.

As an example, consider the following question and the simplified analysis of the responses:

How many pounds do 28 feet of uniform wire weigh, if 154 feet weigh 11 pounds?

A. 2 B. 28/11 C. 11/2 D. 7 E. 14



This question turned out to be a good question: most of the capable students answered correctly while those not as able more often chose the wrong answers. Because it separates students in this way, it was judged as being able to accomplish one of the purposes of the examination, and will appear in the final version.

Because this question is moderately difficult, it would probably appear somewhere near the middle of the paper. If it had proven to be very easy, however, it might have been placed toward the beginning of the final version of the paper with other comparatively easy questions. The theory behind such a format is that if candidates are able to answer the first few questions, they gain confidence and are consequently able to attack the rest of the paper with greater chance of success. It ensures that the maximum number of candidates answer the maximum number of questions.

An even more detailed analysis is carried out on every question which may go into the final edition of the examination, and only those which pass the most rigorous inspection are selected. Despite all these precautions,

however, not every question prepared by the methods described turns out to be a good one. Even after the most careful scrutiny by an expert committee as well as a painstaking experimental screening, some questions may not be relevant, and some may contain errors of fact or interpretation. These, however, are the exceptions.

The Scores: Their Meaning and Use

Each of the candidates' three marks is reported on a standardized scale. The raw score (number of items answered correctly) is converted into a relative measure which has an average of 500 and a standard deviation of 100. This is done for two purposes. First, such scores provide an indication of the individual's relative standing in the group taking the examination, and permit an evaluation in reference to the group. Second, they make possible a comparison of the individual's performance on the different tests. We can more easily answer the question, "Is he better in English, or arithmetic, or equally good in both?"

Scores, and their meanings, may be ascertained from the following figures:

Score	Percentage of applicants with lower score
400	16
500	50
550	69
600	84
625	89
650	93
675	96
700	98
725	99
750	99.4
775	99.7
800	99.9

Scores above 850 signify that the candidate is either one of the most intelligent candidates in the country, or that he has been dishonest. Above 900 there is no longer any doubt that he has had access to the examination or that he has in some way copied the correct answers.

In practice, the candidates with the highest aggregate scores (the total of all three marks) are usually brought in for interviews. Reference to this table can help the headmaster get an idea of how a particular candidate's score compares with others in the country.

Both the order of merit, based on the total of the three marks, and each of the three individual converted scores are reported to participating

schools by the STO.

Editions of the ABC may vary slightly in difficulty from year to year, but these variations do not make any difference to the converted scores.

Some headmasters place applicants in a different order of merit than that provided by the STO. For instance, they may find it more useful to consider only the marks on Paper C and ignore the other two papers; in other words, they may decide to choose only those with the highest aptitude scores for secondary school. If they did this, they would be prepared to give the successful applicants intensive remedial work in subject matter areas when necessary.

Other headmasters might double the Paper B score and add to it the A and C scores, assuming that arithmetic is more important at their schools.

Any participating school may ask the State Testing Office to suggest a way of adding the scores for best results at that particular school. Assistance in this area in one of the many free services the STO can provide.

Preparation for the Examination

The ABC measures abilities that students have developed over a long period of time, so the best way students can prepare for the examination is to read widely and study hard from the beginning of their school careers. The best short-range preparation the pupils can make for the ABC is to read the Candidate Information Booklet carefully. (See the attached copy.) Headmasters are asked to be sure that teachers go through the Candidate Information Booklet with students. Also under the teachers' supervision, students should practice marking a fictitious examination number, do the sample questions, and mark the answers. This will give them an idea of the different kinds of questions they can expect to meet, and how to mark their answers on the answer sheet. The students should then compare their answers with the correct answers which have been marked in the proper manner in the booklet. Then, after they have received permission from their parents to take the examination, they should fill in the application and return the completed forms to their headmasters.

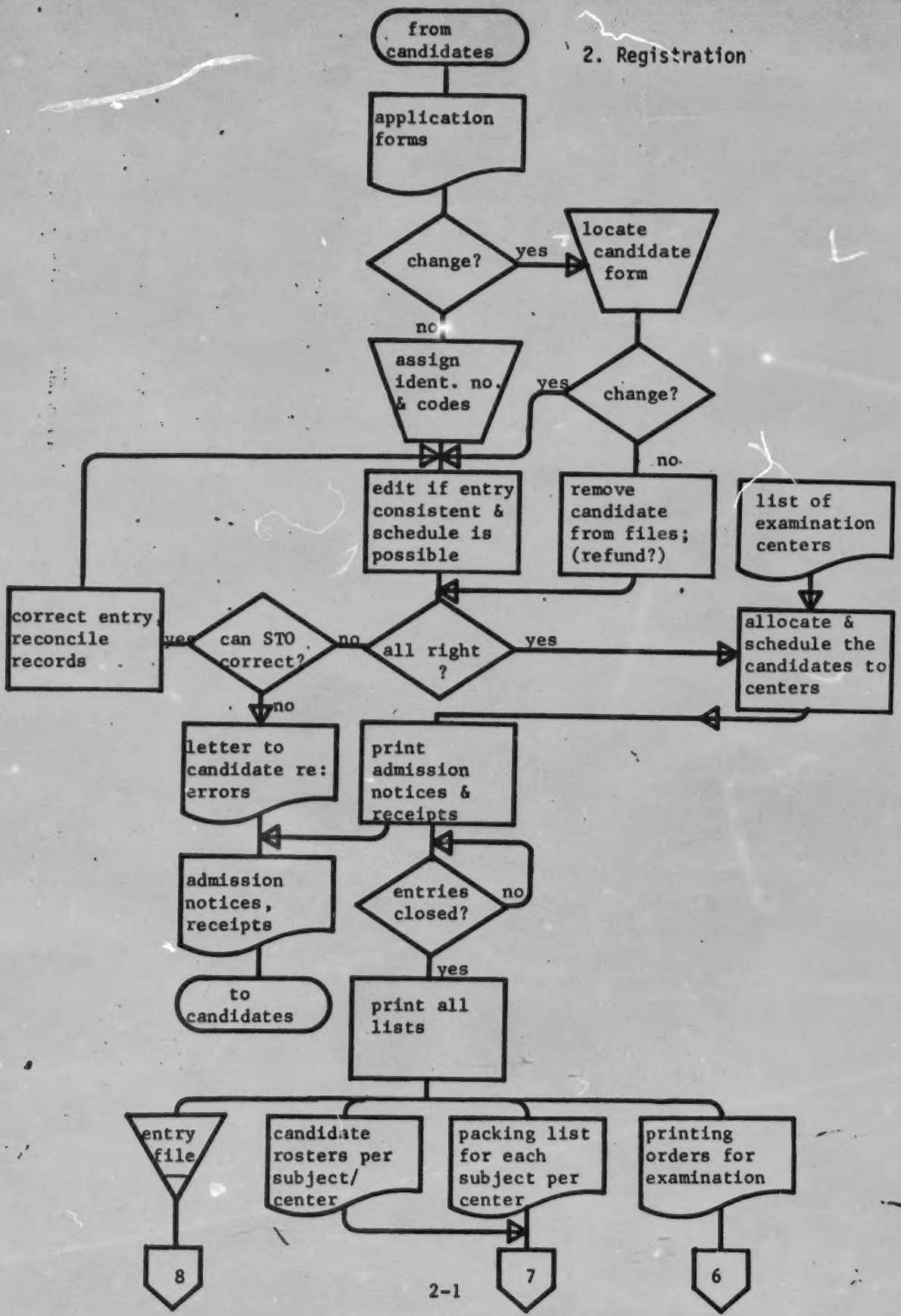
Summary

In the construction of the ABC, the administration of the examination throughout the country, the marking of the papers, the reporting of results, and the offering of consultant services, the State Testing Office makes every effort to help the secondary schools choose those pupils who will

benefit most from further education. Throughout this process, they draw upon the talent of outstanding local university graduates for planning the test, writing the questions and trying them out experimentally, assembling the final edition, and interpreting the results. The STO welcomes suggestions of ways in which it might improve its services. These, or any questions, may be addressed to:

The Director
State Testing Office
P.O. Box 1234
Regina, Atlantis

2. Registration



2-1

REGISTRATION

One of the most important functions of publicity which is sponsored by the testing office is to make clear what requirements an applicant must meet, and exactly how he should apply, to take a specific examination. The candidate information booklet is an excellent vehicle for this information. It also includes an entry form for candidates in the primary schools to complete and give to their headmasters for processing. (The headmaster usually registers candidates in primary school testing programs.) He fills out the roster of applicants from his school, gives pupils' ages, sometimes estimates their ability to succeed in secondary school, and finally certifies that all the entries are full-time students. He may also collect the candidate fees, and forward them to the STO.

At higher levels, the testing office is responsible for supplying an entire book of regulations and application procedures to the candidates through the headmasters, or directly to anyone who requests it. Candidates at these higher levels must complete their own entry forms, which in some cases can be a formidable task. (See Figures 2.1 and 2.2 for examples of the entry form for the WAEC/London GCE. The figures refer to alternate sides of the entry application, which is an 8½" by 11" envelope--a good method, incidentally, of keeping together all papers and other written matter pertaining to a single entry.) To complete these forms, the candidate must have read the book of regulations and studied the timetable because he must be able to arrange a complex combination of subjects. If he can schedule the subjects properly with an acceptable combination of papers, some wags have suggested that he automatically be issued a GCE. When a staff member in West Africa began to flowchart the program to test the legality of every entry, he found that these already considerable difficulties were compounded by a number of errors in the book of regulations--a discovery which further elevates one's opinion of an applicant who completes the application successfully.

A number of certified passport-size photographs are requested (section 13 of Figure 2.2), a precaution made necessary by the large number of impersonations attempted every year. One photograph is kept with the entry application and the other is attached to the admission form, which also doubles as a receipt.

The general flow of work in applications processing is shown on page 2-1. For the more complicated examinations, such as the GCE, the computer is an

excellent tool for checking the legality of each combination of subjects. In addition, it can be used for printing application rejection letters as well as receipts and admission notices.

The candidates may apply to take certain types of examination (e.g., the GCE) at a particular center, whereas they may not indicate a choice for others (e.g., the secondary school selection test) since they may expect to take the examination at their own schools. In either case, the next stage in the process is the allocation of candidates to examination halls. Some testing offices prefer to send candidates from their own schools to others in order to prevent cheating. Other offices consolidate the centers to keep costs down. Any increase in the size of the group which can be handled at a center will reduce the costs. Sometimes this may mean the use of town halls, cafeterias, or churches.

If the testing office keeps a register of approved centers, with information about their seating capacity, physical facilities, and supervisors, candidates can be assigned easily. Without such information, a variety of disasters can occur. In one incident, over 200 candidates were assigned to a center which had facilities for 20; the testing officer had to rent a plane and fly to the centre to dispel the chaos.

Scheduling may be done by hand or by computer, but either way, additional information needed at this stage is the roster of applicants, the packing lists for each center, and the overall printing order for the examination papers. All three can be produced by the computer when one is available. The roster of applicants is used by the administrator to admit candidates into the examination hall and to indicate which ones are absent. The packing lists are based on the number of candidates assigned to a center, plus a margin of extras depending on the size of the group. The lists may include such items as pencils, erasers, and even chalk; an example is shown on page 2-6. The printing order is a grand total of the number of copies of each paper needed for the examination, and would be sent from the data processing room to the reproduction section of the STO as soon as entries are complete.

Sometimes the computer is also used to print the candidate's number on the answer sheet, as shown on page 6-8, or used to punch the school code on an answer card.

A testing office may schedule industrial and civil service examinations weekly at a single center in order to keep testing costs to a minimum.

CERTIFICATE OF EDUCATION EXAMINATION

1970/1

Figure 2.1

2 (a) NAME.....

10. POSTAL ADDRESS.....

.....

11. I have entered for a total of

.....Ordinary level subject(s)

... . Advanced level subject(s) and the fees are as follows :

PHOTOGRAPH

	£	s	d
(a) Basic entry fee of £2			
(b) Fee for.....Ordinary level subject			
(c) Fee for.....Advanced level subjects			
(d) Additional fee for each subject involving practicals at 10/-			
(e) Additional fee for each subject involving orals/dictation at 5/-			
Total			

12. I enclose :

(a) Remittance Letter

(b) Crossed Postal Orders as follows :

Postal Order Number	£	s	d
Total			

13. I also enclose :

(a) Declaration/Signature Slip signed in the presence of a competent witness;

(b) ..... passport size photographs duly certified.

14. I declare that I have studied the Regulations carefully and have made this entry in accordance with them. I accept that I am entirely responsible for the effect of any mistake(s) that I may have made.

15. I agree that the Council may reject my entry or cancel subject(s) if :

(a) not accompanied by the correct fees ;

(b) not in accordance with regulations and therefore irregular;

(c) arrived in the Council's offices after the closing date.

16. I also agree that the WAEC may alter the centre I have chosen for my science practicals.

Candidate's Signature *Date*

2-4

FOR OFFICE USE ONLY

GENERAL ADMINISTRATION SECTION

CASH — AMOUNT FOUND £ — By: — Date: — RECEIPT ISSUED No. — By: — Date:

OTHER ENCLOSURES FOUND — PHOTO No. — By: — Date: — SCIENCE DOCUMENT By: — Date: — DECLARATION By: — Date: — OTHER ACTIONS Nature: — By: — Date:

2. ENTRIES SECTION

FEES CHECKED — PAPERS RECOMMENDED — ENTRY CHECKED By: — Date: — UNDERPAYMENT £ s d By: — Date: — DEFICIT PAID ENTERED By: — Date: — OVERPAYMENT £ s d By: — Date: — REFUND SCHEDULE PREPARED By: — Date: — ENTRIES RATIFIED By: — Date: — OTHER ACTIONS Nature: — By: — Date:

3. EXAMINATIONS ADMINISTRATION SECTION

ADMISSION NOTICES DESPATCHED — ARRANGEMENT FOR SCIENCE PRACTICAL By: — Date: — FOR W.A.E.C. PAPERS By: — Date: — FOR LONDON PAPERS By: — Date: — PHOTO DESPATCHED By: — Date: — OTHER ACTIONS Nature: — By: — Date:

4. ISSUES AND RECORDS SECTION

RESULTS DESPATCHED — FOR W.A.E.C. PAPERS By: — Date: — FOR LONDON PAPERS By: — Date: — CERTIFICATES DESPATCHED — FOR W.A.E.C. PAPERS By: — Date: — FOR LONDON PAPERS By: — Date: — OTHER ACTIONS Nature: — By: — Date:

№ 01859

Figure 2.2

WEST AFRICAN EXAMINATIONS COUNCIL AND UNIVERSITY OF LONDON

GENERAL CERTIFICATE OF EDUCATION EXAMINATION 1970/1

Before completing this form, read all the Regulations carefully. The Council will not accept any responsibility for correcting mistakes made by candidates while completing this form.

1. NAME. Print your name as you wish it to appear on the Certificate. Only one letter, space, or punctuation mark may be placed in each box. 25 boxes are provided – if this is insufficient for your name, you must abbreviate it.

2. NAMES OF CENTRES AT WHICH YOU WISH TO SIT THE EXAMINATION

Written papers .....

Practicals

3. CENTRE NUMBER (Written papers only)

4. OFFICE USE ONLY

5. SEX (Tick as applicable)		6. DATE OF BIRTH			7. FEES REMITTED (See Note 11 over)	8. CORRESPONDENCE COURSE (See Note below)
MALE	FEMALE	DAY	MONTH	YEAR	£ : : d (in words) .....pounds .....shillings.....pence	

9. SUBJECTS. Place X in the box to the RIGHT of the subjects you wish to offer, except where otherwise indicated. Before entering any subjects on this form, check carefully that your choice complies with the Regulations and does not involve any timetable clashes. The notes below show only those rules most frequently broken in previous years. You must study the Regulations and, in particular, the Timetable to check that your choice is valid

UNIVERSITY OF LONDON ORDINARY LEVEL Total of Ordinary Level Subjects →			MARK X
These subjects may be offered with any other in the Ordinary Level syllabus.	ECONOMICS	130	
	ENGLISH LANGUAGE	160	
	GEOGRAPHY	212	
	RELIGIOUS KNOWLEDGE	560	
MATHEMATICS You may offer ONE ONLY of these alternatives	SYLLABUS A		
	PAPERS 1, 2, 3	360.01/02/03	
	PAPERS 1, 2, 4	360.01/02/04	
	PAPERS 1, 3, 4	360.01/03/04	
	PAPERS 2, 3, 4	360.02/03/04	
	SYLLABUS B	361	
	SYLLABUS C	362	
ADDITIONAL MATHEMATICS You may offer ONE ONLY of these alternatives	PAPERS 1, 2	370.01/02	
	PAPERS 1, 3	370.01/03	
	PAPERS 1, 4	370.01/04	
ENGLISH LITERATURE You may offer ONE ONLY of these alternatives	SYLLABUS A	171	
	SYLLABUS B	172	
French may be offered with Music or Navigation. Otherwise choose ONE ONLY of these alternatives	FRENCH	190	
	GREEK	240	
	MUSIC	500	
	NAVIGATION	510	
You may offer ONE ONLY of these alternatives	ITALIAN	330	
	RUSSIAN	570	
	SPANISH	590	
	CLASSICAL HEBREW	718	
HISTORY You may offer ONE ONLY of these alternatives	SYLLABUS A	261	
	SYLLABUS B	262	
	SYLLABUS C	263	
	ANCIENT HISTORY	270	
	BRITISH ECONOMIC	290	
	BRITISH C'WEALTH	280	
You may offer ONE ONLY of these alternatives	BIOLOGY	040	
	BOTANY	060	
Physics with Chemistry may not be offered by candidates offering Physics and/or Chemistry	PHYSICS	540	
	CHEMISTRY	080	
	PHYSICS with CHEMISTRY	550	
Neither of these subjects may be offered by candidates taking subjects 040, 050, 080, 540, 550	GENERAL SCIENCE	200	
	ADDITIONAL GENERAL SCIENCE	205	
ART Candidates for this subjects must offer THREE and ONLY THREE papers At least one paper must be chosen from Papers 1&2 CHECK YOUR ENTRY CAREFULLY		020 01 02 03 04 05	
May not be offered with subject 205	HUMAN BIOLOGY	320	
May not be offered with Latin	GEOLOGY	220	
May not be offered with Chemistry	TECHN. DRAWING	620	
You may offer ONE ONLY of these alternatives.	GERMAN	230	
	LATIN SYLLABUS A	341	
	LATIN SYLLABUS B	342	
May not be offered with classical Hebrew	ARABIC (CLASSICAL)	703	

Note: If you have studied for this examination with a correspondence college operating from U.K. (including overseas branches of United Kingdom based correspondence courses), enter 2 in section 8 above, if elsewhere, enter 1

UNIVERSITY OF LONDON ADVANCED LEVEL Total of Advanced Level Subjects →			MARK X
Before choosing any of these Advanced Level Subjects, study both the Regulations and the timetable with great care. Many of these subjects clash, not only with other Advanced Level subjects, but also with subjects in the Ordinary Level Syllabus	CHEMISTRY	080	
	ENGLISH LITERATURE	172	
	FRENCH	190	
	GEOLOGY	220	
	GERMAN	230	
	GREEK	240	
	LATIN	340	
	MUSIC	500	
	PHYSICS	540	
	RELIGIOUS KNOWLEDGE	560	
	SPANISH	590	
	TECHNICAL DRAWING (ENGINEERING)	620	
	GEOGRAPHY	216	
	ARABIC (CLASSICAL)	703	
	HAUSA	717	
Botany may be offered with Zoology. Otherwise you may offer ONE ONLY of these subjects.	BIOLOGY	040	
	BOTANY	060	
	ZOOLOGY	650	
MATHEMATICS Subject 400 may not be offered by candidates offering subjects 380 and/or 390	PURE	380	
	APPLIED	390	
	PURE AND APPLIED	400	
	FURTHER MATHEMATICS	410	
ART Candidates for this subject must enter THREE and ONLY THREE papers. Either Paper 21 or Paper 22 MUST be chosen. Only one paper from papers 26 and 27 may be chosen. READ REGULATIONS CAREFULLY. CHECK YOUR ENTRY.		020 21 22 23 24 25 26 27	
HISTORY Choose one period in each paper, and write the number of the period chosen under period		260	PERIOD
	PAPER 1 ENGLISH HISTORY	(1-3)	
	PAPER 2 EUROPEAN HISTORY	(1-3)	
	PAPER 3 SPECIAL SUBJECT	(1-11)	
ANCIENT HISTORY Candidate for this subject must enter TWO & ONLY TWO papers. Indicate your choice by an X. May not be offered with HISTORY	See Syllabus for choice of papers	270 01 02 03 04	

WEST AFRICAN EXAMINATIONS COUNCIL SUBJECTS Total of West African Examinations Council Subjects offered at both levels →			MARK X
GOVERNMENT ORDINARY LEVEL		026	
You may offer ONE ONLY of these alternative Subjects	EWE ORDINARY LEVEL	061	
	FANTE ORDINARY LEVEL	071	
	GA ORDINARY LEVEL	081	
	TWI ORDINARY LEVEL	091	
EFIK ORDINARY LEVEL		161	
HAUSA ORDINARY LEVEL		171	
IGBO ORDINARY LEVEL		181	
YORUBA ORDINARY LEVEL		191	
CHEMISTRY ADVANCED LEVEL		144	
ECONOMICS ADVANCED LEVEL		127	
ECONOMIC HISTORY ADVANCED LEVEL		128	
GEOGRAPHY ADVANCED LEVEL		130	
GOVERNMENT ADVANCED LEVEL		126	
MATHEMATICS ADVANCED LEVEL		131	
GENERAL PAPER ADVANCED LEVEL		129	

W.A.E.C. Duplicating

Figure 2.3
Sample Packing List

State Testing Office

. . . . . Certificate Examination

Center Number. . . . . Name of Supervisor. . . . . .

Address. . . . . .

. . . . . .

Received the stationery detailed below: -

Folders		.....	Sealing wax	.....
Loose sheets:	ruled single foolscap paper	.....	Plain envelopes	.....
	graph paper	.....	C.O.S.C. envelopes	.....
	blotting paper	.....	String	.....
Typewriting:	plain paper	.....	Wrapping paper	.....
	flimsy paper	.....	Rubber bands	.....
	carbon paper	.....	Advice notes (#2)	.....
Bookkeeping:	ruled double foolscap paper	.....	Advice notes (#3)	.....
	journal paper	.....	Addressed labels	.....
	cashbook paper	.....	Express labels	.....
	ledger paper	.....	Floor plan	.....
	balance paper	.....	Supervisor form	.....

Logarithm tables .....

General instructions to supervisors .....

Special instructions to supervisors .....

Timetable .....

Certificate of supervision .....

Candidate roster .....

Subject check list .....

Emergency cable forms .....

Please sign and return this form to the Secretary as soon as stationery has been checked.

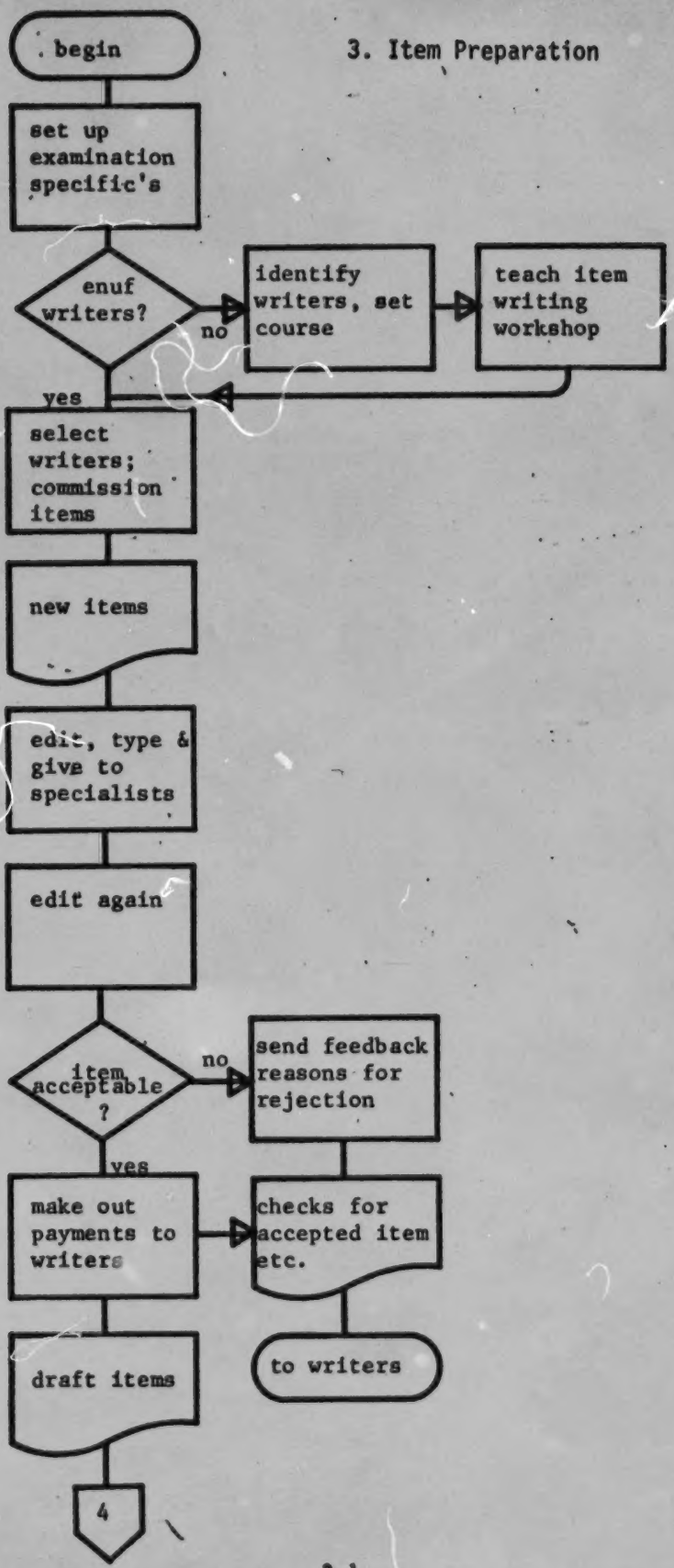
Signature . . . . . .

Date . . . . . .

Important Note: All parcels have been sealed with the official STO sealing tape. If there is any indication that any parcel has been tampered with or if any other sealing has been used, including the official post office sealing tape or paper, take the parcel to the nearest police station immediately, without opening it, and notify the Secretary (telegraphic address: STATUS).

For instance, a general examination may be scheduled to be given at 08.00 every Wednesday at one specific center. No entries are sent to the testing office; instead, applicants for a given position at an industrial plant or a business office are given an admission notice by the personnel officer and told to report to the center. Although no single industry would have sufficient applicants to justify such a schedule, in the aggregate they more than justify it.

Experience leads to one dictum for applicant registration: set up deadlines for acceptance of entries and do not, under any circumstances, change them. If the deadlines become flexible because the mails were held up one year by a flood, the next year by an insurrection in the central area, the following year by a train derailment, the preparation for the examination will not be carried out properly and errors will result. The applicant rosters, packing lists, and printing orders cannot be printed until the entries are closed, so close them once and for all.


begin
3. Item Preparation
set up examination specific's
enuf writers?
no
identify writers, set course
teach item writing workshop
yes
select writers; commission items
new items
edit, type & give to specialists
edit again
item acceptable ?
no
send feedback reasons for rejection
yes
make out payments to writers
checks for accepted item etc.
draft items
to writers
4

ITEM PREPARATION

Before even experienced item writers can write the best items for an examination, someone must have set the educational objectives for each subject. If these are "behavioral" objectives, i.e., if they specify what the student should be able to *do* when he has studied a particular subject, the item writer's job is simplified a hundred-fold. But the major advantages accrue to the teachers and students.

Although the testing office often has little direct responsibility for setting the objectives of education, it can influence whatever curriculum committees produce. And these committees are easily persuaded to move toward more behavioral statements when they see how such a method benefits everyone involved in the educational program, how it improves the quality of the teaching and learning.

Syllabuses usually give little or no direction. Often teachers cannot tell from them what needs to be taught, or what the experts who devised the syllabus thought were important concepts or skills for students to acquire in that subject. The philosophy which underlies behavioral objectives is that good education always changes the learner in some way, and that the change is discernible and demonstrable. When teachers know what changes need to be brought about in order for their students to be "educated" in their subject, they know more exactly what their students should be doing at various points in the process of education, and at its end. Behavioral objectives do not specify how a teacher must teach, what materials he must use, or in what type of situation a learner must learn; they specify only the behaviors which the learner must be able to exhibit at some given point. In this they differ from other objectives which are teacher-centered. Teachers may do whatever they like as long as their students change in the desired direction.

Having clear goals at the outset helps teachers select those procedures and materials which should be most useful to them in helping a pupil reach the goals. And at the end of the course, they can evaluate the effectiveness of their methods and materials more intelligently because of having had clearly in mind from the beginning the changes they wanted to bring about in their charges. They are also better able to make the formal evaluation of candidates which is sometimes required in those countries which take into account the teachers' estimates as well as the official results of the examination.

Clearly stated behavioral objectives channel learning. A pupil can more easily see the purpose for what he is doing, and where it is supposed to be leading him. Also, when a pupil knows where he is headed, he can get there partially or entirely on his own if necessary. He has greater independence and chance for self-direction. In other words, the pupil is helped to learn by the identification in advance of what behaviors he should be exhibiting at various points in his education.

If a pupil knows what he should have learned and be able to do by the end of a course, he is able to evaluate his own progress all the way along. He is more able to see where and when he needs to ask questions and to seek help.

Robert F. Mager, in his book Preparing Instructional Objectives (Palo Alto, California: Fearon Publishers, 1962), gives three rules for writing meaningful behavioral objectives. But first he defines the word "meaningful":

A meaningful objective is one that means the same to all people who read it. If the objective describes the desired final state of the learners in sufficient detail, another teacher, using whatever materials and methods he deems appropriate, will be able to design a course aimed at teaching his pupils to perform in the same desired way. Alternatively, an examiner can write a question aimed more precisely at testing the behavior which the successful learners should be able to demonstrate. In addition, the candidate knows exactly what will be required of him, and wastes little time in the typical student occupation of mind-reading. Thus, rather than being a description or summary of content, a meaningful objective describes an intended *outcome*.

Rule 1. Identify by name what the pupil should be able to do at the end of instruction, the kinds of behavior that will be accepted as evidence that he has achieved the objective.

Here is an objective: To develop a critical understanding of the theme common to "A Sunrise on the Veld" and "Through the Tunnel."

Although this might be an important objective to reach, the statement does not tell what the learner will be doing when he is demonstrating that he is critically understanding the theme. The words which come closest to describing what he should be able to *do* are "critical understanding," and it is doubtful that any two people would agree on the meaning of this term. Because of its vagueness, the teacher's and the item writer's original concept

of it may change from one minute to another without their noticing it.

This objective might be more meaningfully stated in this way: The student should be able to state in his own words the theme common to "A Sunrise on the Veld" and "Through the Tunnel" and support his decision with at least three passages from each story.

The words "state in his own words" and "support" tell what the learner will be doing to show that he has achieved the objective.

So the way to check if an objective meets the first requirement is to be sure it answers the question: What is the student *doing* when he is demonstrating that he has achieved the objective?

Rule 2. Describe the conditions under which the candidate will be required to carry out this activity. For example, what will he be given? And, where appropriate, what will be the time limits?

An objective might thus be stated in this way: At the end of the course, each pupil should be able to play the chorus of "Knees Up, Mother Brown" on a mouth organ within fifteen seconds, with no more than two wrong notes.

This may seem to be reasonably explicit, but someone who was about to take the test would be justified in asking the following questions:

1. Can I use my own mouth organ, or will one be provided?
2. Can I use sheet music, or will I be expected to play from memory?
3. Will I be given a single opportunity to perform, or will I be given a number of tries from which the best can be chosen?

Another sample objective may help to make the point about specifying time limits where they are necessary: To be able to run the 100-yard dash.

If the intention is merely to get the learner to run for a distance of 100 yards without stopping, and if he can meet the expectations of the teacher by so doing, it is not necessary to specify a time limit. But if he must run this distance within 14 seconds in order to pass the requirement, it is only fair to let him know this by stating the objective in this way: To be able to run the 100-yard dash within a period of 14 seconds. The "Knees Up, Mother Brown" objective gives another example of setting time limits.

Rule 3. Specify the criteria for acceptable performance; i.e., what is the pass level?

The final point, after having told the student what you require him to do, and under what conditions, is to inform him how well he must carry out

the activity. This can be done in a number of ways, depending upon the desired end: specifying the minimum number of correct responses which will be acceptable, the number of principles that must be applied or identified, or the number of words that must be spelled correctly; percentage or proportion may be used instead of number. For example: The learner must be able to reply in grammatically correct English to 90 per cent of the English questions put to him during the 15-minute oral examination. Or: Given a map outline of Western Samoa (scale 1:10,000), the learner must be able to mark the locations of Apiă, Lotosamsoni, and Pesega, accurate to within a three-mile radius of their actual locations.

Therefore, three questions can be asked about the objectives to test their clarity and completeness:

1. Does the statement describe what the candidate will be doing when he is showing that he has reached the objective?

2. Does the statement describe the important conditions under which the candidate will be expected to demonstrate his ability?

3. Does the statement indicate how the learner will be evaluated? Does it describe at least the lower limit of acceptable performance?

If the objective is properly stated, the item writer, editing committee, candidates, and teachers will be able to tell when a test question represents the intent described by the objective. They would, for example, recognize that example C below is the only one of the items which is appropriate to test the objective: When asked a question in English, the candidate must be able to demonstrate his understanding of the question by replying in English with an appropriate sentence.

A. Translate the following English sentences.

B. Translate the following English questions.

C. Reply in English to the following questions.

They would also be aware that the following objective is not tested by the question which is applied to it: Given the first three lines of a poem, the learner must be able in 30 minutes to supply a fourth line which conforms to the meter, rhyme, and meaning of the rest of the poem.

Deep asleep, deep asleep, Deep asleep it lies, The still lake of Nyasaland	Insert the number of the best final line from this list of 25 possible last lines. ...

The difficulties raised by a poorly stated objective are manifold for item writers and editing committees, as well as for teachers and pupils. As an illustration, can any of the test items below be considered *inappropriate* for measuring whether a learner has reached the objective: To develop a knowledge of African History?

A. Discuss the meaning of any three significant events in African history.

B. List the names of the leaders of revolutions in Kenya, the Congo, Nigeria, Lesotho, Ghana, Sierra Leone, Dahomey, Togo, and Uganda.

C. List as many events as you can that occurred in African history between 1850 and 1950 and give the date of each.

How must direction does this objective give any of the people concerned?

This last objective provides an excellent illustration of what words *not* to use in writing objectives. Certain words and expressions commonly used in syllabuses are so vague and open to so many interpretations that they give no help whatsoever to anyone, and succeed in rendering the objective useless. Some of these which Mager and others list are:

know	have faith in
understand	have a feeling for
really understand	comprehend
appreciate	become aware of
fully appreciate	develop a knowledge of
grasp the significance of	become familiar with
enjoy	create an awareness of
believe	have a critical understanding of
work with	deal more effectively with

In contrast, some committees use words and expressions open to fewer interpretations and therefore of more value. Listed in generally increasing complexity of the thinking they require, they are:

recall	select	suggest	manipulate apparatus
list	make	derive	recognize & cite evidence for
choose	do	expand	prove
find	reproduce	interpret	organize data
gather data	compute	summarize	analyze
investigate	measure	trace	express
describe	estimate	convert	apply
define	prepare	translate	distinguish between
finish	build	classify	construct
complete	use	illustrate	devise a method
indicate	recognize	demonstrate	plot a graph
label	examine	write	state a problem
locate	identify	recite	identify the variables
match	put in order	participate	solve

contrast	discuss critically
compare	predict
specify limitations and assumptions	design
organize	deduce
relate	integrate
differentiate	produce
reformulate	propose reasons and defend them
justify	formulate hypotheses
estimate	reorganize
interpret	discover
generalize from data	manipulate ideas
synthesize	evaluate
infer	create
debate	

Comparison of a section of the geography syllabus in Malawi before and after revision by the committee shows a vast improvement in clarity and specificity which should lead to better instruction and better examinations:

BEFORE- The teacher should cover the following topics:

A. Air Pressure and Winds. Movement of air: wind systems of the world. Measurement of pressure and wind.

B. Wind and Precipitation. Relation of rain to moisture bearing winds.

C. Climate and Vegetation. Use of photographs to relate vegetation to climate.

AFTER- At the conclusion of the course the student should be able to:

A. Define air pressure and explain how winds are produced.

Given a wind rose, state the frequency, force, and direction of the wind.

Write out the Beaufort Wind Scale.

Define, and use correctly in context the following terms: windward, leeward, windbreak, on shore and off shore winds, turbulent and laminar air flow, stability and instability, air mass, front, Chinook.

Recognize from a photograph or drawing and describe the weather conditions associated with the following cloud types: nimbus, cumulus, stratus, as normally associated with the height names cirro and alto

Use or explain how the following meteorological instruments are used: weather vane, Six's thermometer, wet and dry bulb thermometer, rain gauge.

Describe weather conditions for a station for which the weather elements are shown in symbols or be able to interpret a rainfall scatter diagram, evaluate means and averages, and complete an Isopleth map.

B. Describe the three main types of rainfall and give examples of when and where they are likely to occur.

List the factors of climate and clearly differentiate between weather and climate.

Describe the main features of the climates of those areas covered regionally.

C. Describe and recognize from photographs the major vegetation types: forest, savanna woodland, scrub and grassland, and attempt to relate the factors controlling the type of vegetation. (Greater emphasis will be placed on accurate description of the vegetation and the assessment of factors revealed in the photograph than in determining the location of the area or its climate.)

Assuming for the moment that thought processes can be regarded as falling into two categories--high and low level, experience has shown time and again that when the committee writes behavioral objectives for the syllabus, more statements are produced which call for high-level thinking than with any other method of preparing educational specifications. The reason may be that the demand for increased specificity forces them to take a closer look at their goals, seriously rethink and clarify them. Most syllabus committees are eager for students to learn to think, not merely to memorize, and their behavioral objectives are able to reflect this attitude. In his turn, the item writer is also affected. He must thereafter devise questions which test the pupils' ability to think in certain ways, not only their ability to recall. Thus the entire educational process increases in quality.

In contrast, when course content alone is presented, the temptation for both teachers and item writers is to concentrate on low-level thinking and behaviors since no other goals are specified. When the traditional blueprint is used (see page 3-16), only if it requires a certain number of high-level test items will any reasonable quantity of them be produced at all.

The terms "high" and "low" level thinking need some clarification. In the mid-1950's, a committee of college and university examiners in the United States devised a classification of thought processes. One important objective of their work was to get teachers and examiners to ask higher level questions and thus to demand that students use more of their brain power and creativity. Essentially, their six classes can be reduced to two: high level thinking and low level thinking. In the first category fall all questions of the types described under levels two through six below; in the second, all questions of the type described under level one. What follows is a condensation of this committee's now-famous classification.

The material[1] is presented in some detail for two reasons: (1) study of the classification by testing officers and syllabus committee members often will lead to the setting of improved test questions; and (2) many testing offices do not have ready access to the Bloom volume in spite of its usefulness.

Level 1. Knowledge: of facts, of conventions, of methodology, of terminology, of principles and generalizations, etc. Anything that can be memorized and recalled.

Examples of questions which call for this level of thinking:

1. Who wrote the poem "A Telephone Conversation"?

   A. Leopold Senghor
   B. John Pepper Clark
   C. Wole Soyinka
   D. Christopher Okigbo

2. For computational purposes, forces are frequently represented by

   A. straight lines.
   B. circles.
   C. arcs of a circle.
   D. angles.
   E. objects of three dimensions.

Level 2. Comprehension: This is the lowest level of understanding. The emphasis is upon the grasp of the meaning and intent of the material. There are three sublevels, of ascending difficulty, within this second level:

A. Translation: The ability to understand metaphor, symbolism, irony, exaggeration; ability to read music, maps, mathematical symbols, architectural plans; ability to condense and paraphrase; ability to give examples to explain something abstract; ability to translate one language into another.

Sample questions are:

1. The poet's main point in "Piano and Drums" is that

   A. he longs to return to the innocence and simplicity of childhood.
   B. he longs to go to far away countries and live a more complex life.
   C. he is caught between two opposing ways of life.
   D. the world is on the brink of a new age in which anything could happen.

2. When a current is induced by the relative motion of a conductor and a magnetic field, the direction of the induced current is such as to establish a magnetic field that opposes the motion. This principle is illustrated by

---

[1] Bloom, Benjamin S. (ed.). A Taxonomy of Educational Objectives. New York: David McKay Company, 1956, 206 pages.

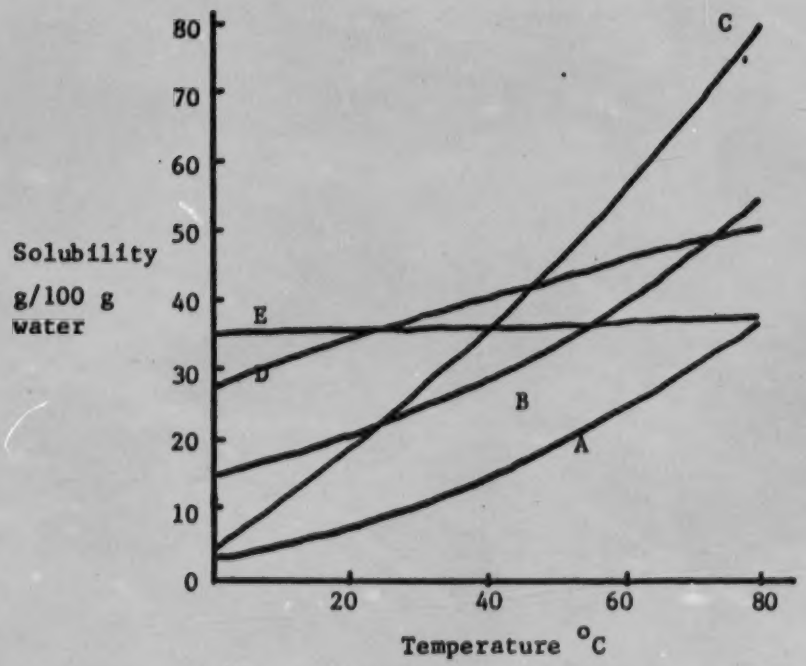
A. a magnet attracting a nail.
B. an electric generator or dynamo.
C. the motion of a compass needle.
D. an electric doorbell.

B. Interpretation: A person's ability to grasp the thought of the work as a whole, to get some total view of what the communication contains, and to relate it to his own fund of experiences and ideas. (Translation takes part by part; interpretation deals with the whole. Notice that each of these levels is just one step higher and requires harder thinking than the last one. Each one is also dependent upon the person's ability to reach the last level.)

Sample questions are:

1. Which of the following is the best title for the passage?

A. Types of Drumming
B. The Living Tradition of Drumming
C. Fees for Drumming Lessons
D. The Head of the Drummer



2. Which saturated solution contains the lowest concentration of dissolved solid at 18°?

C. Extrapolation: The ability to extend trends beyond the given time, topic, or sample.

Sample questions are:

1. In "A Sunrise on the Veld," when the boy goes off by himself alone to think the next day, which of the following conclusions is he most likely to arrive at?

   A. He will decide never to hunt again because how he realizes that animals have as much right to live as he does.
   B. He will decide never to hunt carelessly again since he has probably been responsible for causing many wounded animals to be eaten by ants.
   C. He will decide to tell his parents where he is going from now on in case he should hurt himself and be helpless like the buck.
   D. He will decide to stop the natives from throwing stones at animals by reporting the incident of the buck to the local authorities.

2. Which salt would be most suitable for growing crystals by cooling a hot solution? (Based on diagram above.)

Level 3. Application: The ability to use abstractions, generalizations, techniques in new situations. The ability to carry over one's knowledge into an area in which he has not previously applied it. This carries extrapolation one step further.

Sample questions are:

1. In The River Between, why is it ironic that the white missionaries forbade circumcision?

   A. Because there is nothing evil or physically harmful about circumcision. Many doctors of all faiths now recommend it for all infant boys.
   B. Because the missionaries themselves were probably circumcised.
   C. Because Jesus and all the first Christians were circumcised.
   D. Because later in the book they recognize their mistake and begin to encourage the practice.

2. An electric iron (230 volts, 1000 watts) has been used for some time and the plug contacts have become burned, thus introducing additional resistance. How will this affect the amount of heat which the iron produces?

   A. The iron will produce more heat than when new.
   B. The iron will produce the same heat as when new.
   C. The iron will produce less heat than when new.

Level 4. Analysis: This includes three distinct sublevels;

A. The ability to break down a communication into its parts, including;

   1. to recognize unstated assumptions.
   2. to distinguish facts from hypotheses.
   3. to identify motives.
   4. to identify conclusions and supporting statements.

B. The ability to make the relationships among the parts clear including abilities:

1. to identify which facts or assumptions are essential to the main argument.
2. to distinguish between major and minor themes.
3. to distinguish cause and effect relationships.

C. The ability to recognize the organizational principles including:

1. to recognize form and pattern in literary and artistic works as a means to understanding their meaning.
2. to be able to infer the author's purpose, point of view, etc.
3. to see the techniques used in advertising, propaganda, or in any persuasive materials.

Sample questions are:

1. Why doesn't the boy care about playing with the bigger boys any more after he swims through the tunnel?

   A. Because he no longer needs to prove himself.
   B. What they are doing then seemed like child's play compared with swimming through the tunnel.
   C. Because he knows they will only rebuff him again if he tries to play with them.
   D. Because he is too emotionally and physically exhausted.

2. Which of the following assumptions is necessary in order to determine the mass of a drop by the method described?

   A. The drop falls with uniform acceleration.
   B. All the drops sprayed into the chamber are of the same size.
   C. The drop is nearly spherical.
   D. The drop is charged.
   E. The electrical force is equal to the gravitational force.

Level 5. Synthesis: The ability to put together elements and parts in such a way that they make a pattern or structure not clearly present before. This category covers most creative behavior. The emphasis is upon uniqueness and originality. This level differs from the others in that the end product is always *more* than the material given to work with.

Not all poems, short stories, plays, paintings, musical compositions, etc., are creative or demonstrate synthesis; that is, creativity is not dependent upon form. It can and does occur anywhere, in any field of study or action, at any time.

> . . . Although literature, fine arts, music, and drama seem to be the most popular media, creative expression need not be limited to these. Such activities represent synthetic processes to the extent that they require the individual to organize ideas into new patterns, and

probably many of them do. However, many do not qualify because they emphasize expression of emotional impulses and physical movements, rather than organization of ideas (Bloom, p. 165.).

Level 6. Evaluation: The ability to make judgements about the value of ideas, works, solutions, methods, materials, etc., on the basis of clear criteria. It is this matter of criteria which makes this level higher and different from all the others.

Levels five and six are extremely difficult to use in large-scale examinations since the emphasis on uniqueness of responses precludes machine marking at the present state of the art. In addition, formulating multiple-choice questions at these levels is probably beyond the powers of the item writers presently plying their trade. However, a few illustrative essay questions are presented.

Sample questions for level five are:

1. Plan a unit of instruction for No Longer at Ease.
2. Outline a set of headings that might be suitable for any autobiography or biography.
3. Give as many possible reasons as you can why politeness has come to be considered a prime virtue in Malawian culture.
4. You have been awakened in the night when a six-legged, hard-shelled, cold-blooded creature, a foot and a half across, crawled over your feet. Name it and describe its possible life styles.

Sample questions for level six are:

1. In Mr. Stuart's argument about the importance of a "great past," to what extent do his conclusions follow logically from the material presented?
2. What criteria are appropriate for judging an improvised drama?
3. What was the effect of Galileo's work with the telescope?

Well-stated objectives which call for some complex thinking give the item writer the direction he needs as well as ensure that test items will be of higher quality. It is not the item writer's responsibility to define the educational goals, although this happens far too often simply because no one else does the job. This is the way examinations come to determine the syllabus; termed the "backwash effect," this is understandably a most undesirable state of affairs.

To simplify matters for the item writer further, the group responsible for setting up the model for the examination should note how much emphasis they wish to place on each objective. The proportion of items needed on each topic or objective could be dictated by percentage weights, which might

be assigned as in the following example:

1. Given an English sentence, a candidate should be able to express the statements algebraically. The expression should contain no more than three unknowns. Any symbols used must be clearly defined. (7%)

2. For a statement involving no more than two quantities, he should be able to express the information in a suitable graphical form. The algebraic expression may be linear or non-linear; if non-linear, it will be of a form covered by those mentioned under "functions and their graphs." (8%)

3. A candidate who has been given or led to discover a simple relationship between a number of variables should be able to express the relationship in the form of a formula. (5%)

4. Given a general formula involving p ($p<5$) variables, the candidate should be able to interpret the formula in a particular situation. Given the values of p-1 of the variables, he should be able to determine the value of the remaining one. (3%)

5. If a variable x is expressed in terms of variables p,q,r,s, in an algebraic form which is a polynomial, rational form, or involves square or cube roots, then the candidate should be able to re-arrange the formula to express p in terms of x,q,r, and s. (0%)

6. Given any . . . .

Notice that some objectives may not be emphasized this year (0% for objective six), but next year the percentages could easily be different. Consequently, teachers try to help every pupil demonstrate all the behaviors in the syllabus every year, the syllabus gets general coverage, and the pupils are more likely to receive a better education.

In some instances the committee may not assign different weights to the objectives but may request that the syllabus be "covered" in the examination. For this situation the tester takes a representative sample of objectives from the total curriculum for each examination and writes questions based on them. Figuratively, he places a statement of each objective in a box, draws out one at a time, writes a number of questions based on it, then draws out another, etc., continuuing until he has the number of questions he needs. He would probably not draw out every objective. The next time an examination must be set, he repeats the entire procedure, which means that the same objectives may, or may not, be tested.

But if the syllabus does not specify any behavioral objectives, and only course content is listed to guide the item preparation, the testing office is largely on its own. Either they or a committee must draft a blueprint for the examination. This document is basically an orderly arrangement of examination material on two dimensions. One dimension refers to the subject

matter--the content of the syllabus generally, and the other refers to the level of thinking entailed. The entries in the cells denote the relative emphasis to be placed on the particular combination of topic and level of thinking. An example of a weighted blueprint appears in Figure 3.1.

Some testing offices use more than the two categories suggested, but since a primary aim has always been to get more high level and fewer low level questions into the examinations, two categories have always proved adequate. Although some offices think that increased complexity in the categorization will lead to a noticeable improvement in the examinations, this has not generally proved to be the case.

No matter how the testing office is set up, or how well the syllabus and testing committees work together to prevent it, the examinations have an effect on the actual curriculum. So even when the syllabus committees have specified behavioral objectives, and the "backwash effect" of the examinations is thus diminished, the office must examine and re-examine its tests for any traces of poorly-conceived questions which would lower the quality of the curriculum.

After the specifications are determined, whether through the use of behavioral objectives or the traditional blueprint, a variety of decisions must be made before the actual commissioning or writing of the items: the particular form the test will take, the general item types, the reference materials needed or permitted for each section, the total number of questions, the level and range of difficulty of the questions, and so on. These questions settled, writing may begin.

Seldom is a single examiner able to produce the quantity of items needed for large-scale examination programs. (For instance, more than 2,000 items were needed for just one examination in Malawi.) To complicate matters, no large-scale examination is sufficiently secure to allow extensive re-use of the items. Therefore, item writers must be trained, and since it is much easier to teach item writing to subject matter specialists than to teach the subject matter to item writing specialists, the most likely group to start with is classroom teachers. Often the ministries of education locate about fifteen to thirty experienced teachers, and arrange to have them boarded at some central location for a one-week item writing course.

On the first day, a member of the STO staff introduces the topic of objective testing, explains the principles of item writing, and demonstrates

Figure 3.1 Example of a weighted blueprint.

		Unit A						Unit B			
	Content→ / Skills↓	Change, functions, graphs	Increments, gradients	Derived functions	Differential equations	Composite functions	Inverse functions	Area by lim. process	Definite integrals	Area - any curve	Log. function, base e
lower order	1. Mathematical vocabulary/symbolism		2	2		1	3				
	2. Facts and generalizations	1								2	1
	3. Calculate/compute		2	3	2	2	1		1	2	1
	4. Common patterns of argument	1	1	2			3				2
	5. Analyze/translate	2	4		2						
higher order	6. Logical consequences of mathematical results				2					1	2
	7. Conclusions							1			
	8. Mathematical models				2						

the steps involved in general test construction. One or more members of the syllabus committees then explain the detailed requirements for a particular examination.

During the remaining days, the participants write items, submitting each one on a single sheet of paper (shown in Figure 3.2). As each one is received, the staff member and syllabus committee member edit it, and either keep it if it is acceptable, or return it to the writer with suggestions for revision. A collection of materials used to teach this course is included at the end of this section (page 3-21).

After the course is over, the items are reread and separated into those to be trial tested, and those presently unusable. The former are edited carefully, and then subjected to the following interrogation:

1. Is the item as clear as possible?
2. Is it as precise as possible?
3. Are the necessary qualifications included?
4. Is only the essential specificity given?
5. Is the difficulty level appropriate?
6. Are irrelevant clues lacking?
7. Is a concensus possible on the "best" answer?
8. Are all responses appropriate to the stem?
9. Are all distracters plausible and attractive?
10. Are the responses non-overlapping?
11. Are the responses in logical order, or mixed position?

If the answer to any question is no, the item is edited again, or discarded if it is beyond hope.

The costs of the item writing course vary considerably, depending on the physical accommodations and the time of year during which it takes place. If the course is scheduled for a holiday period, it is often possible to board the participants at a school and keep the costs down. Of course when participants are deprived of their normal holiday, they are paid for the items they write, as detailed below. When the ministry prefers to schedule the course during a school term, there are additional costs for boarding but no pay for the items.

All participants of the course who have had at least one item accepted are considered when the time comes for the commissioning of items. The examiner sends each one specifications concerning the number and type of

Figure 3.2 Instructions to item writers for submitting items, and submission form.

1. The name of the subject, and behavioral objective specifications should be filled in on the form in the spaces provided.
2. The item should be typed or written clearly, with a separate sheet being used for each item. If supplementary material is being used which cannot be fitted into the space provided, attach it securely to the relevant sheet. The source of all supplementary material should be clearly stated.
3. In all items the correct option should be written first and the letter A underlined.
4. The estimated difficulty should be filled in. In assessing the difficulty of items, which should always be related to the level of the candidates for whom the examination is designed, the categories easy/average/difficult should be marked as appropriate. Definitions are as follows:
   Easy: an item which is correctly answered by more than 70 per cent of the candidates.
   Average: an item which is correctly answered by 30 to 70 per cent.
   Difficult: an item which is correctly answered by fewer than 30 per cent.
5. If an item can appear under more than one behavioral objective, this should be noted when filling in the specification.
6. When the item is ready to be returned, the writer should sign and date the declaration.

.....

(on reverse side)

Subject: . . . . . Specification:. . . . . .

Estimated difficulty: ( ) easy Alternate spec: . . . . . .
( ) average
( ) difficult

Item:

Declaration: Under penalties of perjury, I declare that I have written the above item and have communicated neither the substance nor nature of the item to anyone. To the best of my knowledge the item is secure, and I promise not to divulge its substance or nature in the future.

Dated: _____ Signed: _____

items needed, and the necessary reference materials. In general, four times as many items are commissioned as will be needed in the final draft. A supply of the forms shown in Figure 3.2 is sent for the item writer's use.

When the items are received at the STO, the packet is checked for the presence of the table of specifications and any other materials which were sent to the writer. If all materials have been returned, and the declaration has been signed for every item, a receipt is sent to the writer.

Then an officer at the testing office goes through the items to make them technically acceptable. If the officer is competent in the subject, he checks the content of each item. In particular, he makes certain that there is one and only one correct answer to each item and that the distracters are plausible.

The items are then typed out double-spaced for reproduction, with the reference number above each question, and the originals placed in security storage; see chapter nine. After proofreading the copy, the officer has them duplicated and a copy sent to each member of the subject committee. Each officer in the STO is also given a copy and asked to make his comments on the items.

When the subject committee members and STO staff meet, they discuss the various items, possible amendments, and rejections. One member records all alterations made and the reasons for rejecting the item, if necessary. Payment is made for items which successfully pass this stage. A rough idea of the amount to be paid per item can be found by dividing a typical professional's daily salary in that particular country by ten, since an experienced item writer ordinarily produces about ten good achievement items per day. In general, questions for which no payment is made are: (1) rejected items; (2) those which do not have a correct answer; and (3) items in which more than two options have been amended. Some offices give the writer an additional payment if the item passes the pretesting stage, a practice which makes the writer happy but which does not contribute to the production of better items. In general, better items are produced in direct proportion to the amount of feedback given the writers; thus it is wise to send rejected items back with comments concerning possible ways of improvement. If the item was rejected at the pretesting stage, item analysis information can also be sent.

If the items cannot be considered by a subject committee, the use of

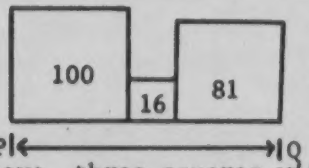
extremely competent subject specialists for the final editing before pretesting is recommended. No amount of item analysis will compensate for the critical eye of the specialist at this stage. In fact, sometimes it is both necessary and advantageous to go outside the syllabus or examining subject committee for this critical and unjaundiced eye.

All items which have been accepted are placed on cards, as shown on page 5-4 and the draft submission forms are destroyed by fire. These cards, which make up the item file, include information concerning the author, category of question, changes made in wording, source of material, etc. After a few years of testing office experience, items can be drawn from the file when additional ones are needed in later stages of the testing cycle. Some offices even repeat items after a few years, but never in a set sequence; (i.e., the same examination or even one part of it is not repeated every seventh year). If the items are higher-level ones based on a set of behavioral objectives, the pupils' study of the old test papers is equivalent to study of the textbooks, and therefore not objectionable.

When the items are transferred to the cards, and editing is complete, the stage is set for pretesting them on live candidates.

Some Notes on Item Writing

The item (sometimes called the question) is made up of the stem, which is an introductory question or an incomplete statement, and two or more responses, which are suggested answers to the question or completions of the statement. The responses are called alternatives, and consist of the correct answer and some incorrect answers termed distractors. For example:



In the figure above, three squares with areas of 100, 16, and 81 lie side by side as shown. By how much must the area of the middle square be reduced in order that the total length PQ of the resulting three squares be 21? } Stem

A. $\sqrt{2}$
B. 2
C. 4
D. 8 } Distractors
E 12 Key, or correct answer } Alternatives (or options)

Some very useful stems for achievement tests are:[1]

1. Purpose
   - A. What purpose is served by . . .?
   - B. What principle is demonstrated by . . .?
   - C. Why is (this) done . . .?
   - D. What is the most important reason for . . .?

2. Cause
   - A. What is the cause of . . .?
   - B. Under which of the following conditions is . . .(this) true?

3. Effect
   - A. What is the effect of . . .?
   - B. If (this) . . . is done, what will happen?
   - C. Which of the following should be done . . .(to achieve a given purpose)?

4. Association
   - A. What tends to go with . . .?
   - B. How often do A and B go together?
   - C. If we do . . . to A, what happens to B?

5. Recognition of Error
   - A. Which of the following is wrong (in the given statement)?

6. Identification of Error
   - A. What kind of error is this: . . .?
   - B. What is the name of this error: . . .?
   - C. What recognized principle is violated when . . .?

---

[1] Adkins, Dorothy C. Construction and Analysis of Achievement Tests. Washington: Government Printing Office, 1947, out of print.

7. Evaluation
   A. What is the best evaluation of . . .(for a given purpose and for what reason)?

8. Difference
   A. What is the (or an) important difference between . . .?

9. Similarity
   A. What is the (or an) important similarity between . . .?

10. Arrangement
   A. In the proper order (to achieve a given purpose or to follow a given rule), which of the following comes first (or last, or follows a given item)?

11. Incomplete Arrangement
   A. To produce the proper order, which of the following should be inserted (here) to complete the series?

12. Common Principle
   A. What is the principle . . .?
   B. Which item does not belong . . .?
   C. Which of the following items should be substituted . . .?

13. Controversial Subjects
   A. Although not everyone agrees on the desirability of . . ., those who support it do so primarily for the reason . . .?

In the construction of test questions, certain principles and precautions are adopted in order to eliminate technical weaknesses. These are classified below as they refer to (1) the question as a whole; (2) the stem of the question; and (3) the answer alternatives.

The question as a whole

1. Each question should test a concept or idea that is important for the examinee to know. Avoide unnecessary technical terms or minute obscure details. In other words, one would *not* ask such a question as this:

> In whose house were Dr. Johnson and Boswell staying when Johnson said, "This woman would sink a ninety-gun ship. She is so dull--so heavy."
> A. Sir Alexander's
> B. Mr. Maclaurin's
> C. Lord Lochbuie's
> D. Mr. MacLeod's

2. Each response should require the examinee to give a reasoned response rather than fragmentary factual information.

3. Vary your questions in difficulty. The good question in one which 50 to 60 per cent of the examinees can answer, but there should be items which can be answered by as few as 15 per cent or as many as 85 per cent. Most beginners tend to underestimate the difficulty of their questions.

4. Use clear and simple language, free of ambiguity and appropriate to

the educational level of the examinees. The language should be overly-simple English so that the questions measure knowledge of the subject matter and not reading ability.

5. Avoid excessive window-dressing and material which is irrelevant to the purpose of the question. Attempt to reduce the reading time and complexity. If a long presentation is absolutely necessary, write more than one item on the information so the candidate receives a reasonable return for his work. The following example illustrates the very common habit of introducing non-essentials:

> Psychological tests have many important characteristics. Most competent authorities would agree, however, that the most essential of the following characteristics is... etc.

6 6. If the purpose of the question is to test ability to apply knowledge or to think critically, make sure that the item cannot be answered on the basis of factual knowledge alone.

7. Avoid negative questions whenever possible and double negatives always. One of the most baffling double negatives ever seen goes like this:

> Albany is not the capital of New York State. True or False?

8. Avoid trick questions. Even if they seem too clever to be resisted, remember that it is usually the good candidate who gets misled by them. An example of the type to resist is:

> The Battle of Hastings was fought in
>
> A. 1066 B.C.
> B. jest.
> C. word and deed.
> D. Sussex.
> E. all of these.

<u>The stem of the question</u>

1. Develop the question around some central idea or problem which is clearly stated in the *stem* and to which all the options refer. The function of the stem is to set the stage for the alternatives which follow.

2. Give the stem either as an incomplete sentence or as a question. The choice should be made on the basis of which seems to be best for the particular test question. Some research evidence suggests that the question format is easier for younger candidates and the incomplete sentence for older ones.

3. In the stem of the question, make clear to the examinee the nature of the answer that is desired. He should not be required to reread the stem after inspecting the alternatives. The candidate who knows the answer should be able

to state it after reading the stem and *before* reading the alternatives.

4. In general, include in the stem any word that must otherwise be repeated in each alternative.

5. Avoid stereotyped phraseology which places too high a premium on rote memory and gives credit to examinees who lack a basic understanding of concepts. For example:

A test is valid
- A. to the extent to which it measures what it is supposed to measure.
- B. if a person gets the same score every time he takes it.
- C. if 50 per cent of the examinees score above the mean.

6. Examinees who are accustomed to selecting the one correct response are confused when suddenly required to select the one incorrect response. If a negative aspect is included, underline, italicize, or use upper case letters to emphasize the appropriate words.

The answer alternatives

1. The alternatives in an item are as important as the statement of the problem in the stem. They should be chosen carefully--incorrectness alone should not be the criterion. The following are frequently the sources of good foils:

a. a common misconception.

b. a statement which in itself is true but does not satisfy the requirements of the problem.

c. a statement which is either too broad or too narrow for the requirements of the problem.

d. a "nonsense" response which sounds plausible to the uninformed.

2. The difficulty of a question depends largely on the alternatives. The finer the distinctions that must be made to select the correct alternative, the more difficult the question.

3. Make all the options similar to one another in the kind of relationship they have to the thought in the stem:

The Yearling
- A. is a story of the Ozark foothills.
- B. describes life in the Washington timber country.
- C. has as its setting the Allegheny Mountains.
- D. has the Florida scrub as its locale.

This item would be improved if the stem were stated thus:

The Yearling takes place in

4. Each alternative should be grammatically consistent with the stem, *not* like the following examples:

Among the causes of the American Civil War were

A. Southern jealousy of Northern prosperity.
B. Southern anger at interference with the foreign slave trade.
C. Northern opposition to bring in California as a slave state.
D. differing views on the tariff and constitution.

Calcium is an

A. element.
B. mixture.
C. compound.
D. None of these.

5. Avoid the telegraphic style; that is, the tendency to omit words such as "a," "an," "the," "and."

6. Make the alternatives independent and mutually exclusive. Overlapping options allow the unprepared examinee easy credit. In addition, if two of the alternatives are opposites, one is most likely to be the correct answer. Following is a question for which all the answers are correct:

In 1970 the population of the U.S. was

A. over 100 million.
B. over 125 million.
C. over 150 million.
D. over 175 million.
E. over 200 million.

7. Guard against irrelevant clues in the correct answer, such as the use of textbook wording in the correct answer but not in the other alternatives, and rote verbal associations between the stem and the correct answer.

8. Avoid the general tendency to make the correct answer consistently longer than the other alternatives. This tendency results from the writer's desire to qualify his correct answer so that it will be completely and exactly right, while the other options are set down as flat or simple statements. The following example illustrates this fault:

Why were the Republicans ready to go to war with England in 1812?

A. They wished to honor the American alliance with France.
B. They wanted additional territory for agricultural expansion and felt that such a war might afford a good opportunity to annex Canada.
C. They were opposed to Washington's policy of neutrality.
D. They represented commercial interests which favored war.

9. Avoid giveaways in the alternatives such as "always," "only," "every," "all," "never." The question writer should ordinarily test a more sophisticated

concept than recognition by the examinee that statements are rarely universally true. Such words may, however, be advantageously used on occasion, particularly if some questions are also designed in which the "always" or "only" response is the correct one.

10. When "none of these" is used as the response to several questions, be careful that it is not always used as the correct answer. This also applies to "all of the above."

Suggestions for Preparing Essay Tests

1. Restrict the use of the essay examination to those situations to which it is best adapted. When it is not clear that the essay is required for measuring the desired objective, use the objective format.

2. Ask questions or set tasks which will require the examinee to demonstrate his command of essential knowledge.

3. Ask questions which are determinate in the sense that experts could readily agree that one answer is better than another.

4. Generally give preference to more specific questions that can be answered briefly.

5. Increase the number of questions asked and reduce the amount of discussion required for each. Always indicate clearly the type of discussion desired.

6. Avoid giving examinees a choice among optional questions, because:
   a. studies have shown that examinees are not always, or even usually, able to choose the question on which they will perform the best;
   b. choice complicates measurement and introduces factors of judgment extraneous to the factors measured. For example, presumably you are not measuring how well the examinee can choose a subject;
   c. this method helps the poorly prepared examinee but not the well prepared one. Thus you discriminate against good examinees.

7. Instead of giving a choice, increase the number of short questions and require them all to be answered.

8. Test the question to see if it is possible to write an answer to it by trying to do the task yourself.

Sample essay questions for science[1]

1. Comparison and contrast

   A. What are the differences between weather and climate?
   B. In parallel columns compare the eye and the camera, noting all similarities and differences.

2. Decision for or against

   A. Will air serve as well as oil in a hydraulic press? Explain.
   B. Which system of measurement is more logical to use if you have learned both equally well? Justify your answer.

3. Application of facts or principles already learned to new situations

   A. Suggest ways of correcting annoying reverberations in a hall or church.
   B. A glass stopper is stuck in a bottle and cannot be removed by twisting or tapping it. Suggest a way of loosening it.

4. Classification

   A. To what class of lever does the wheelbarrow belong?
   B. From your observation of the number of legs this specimen has, to what group of animals do you think it belongs?

5. Relationships involving cause and effect

   A. What is the relation between efficiency and friction in a machine?
   B. What would be the result of washing a piece of cake down your throat without chewing it?

6. Example or illustration

   A. Name a parasite.
   B. Describe a case that you have actually witnessed where friction proved to be a disadvantage.

7. Statement of author's aim or purpose in the selection or arrangement of material

   A. Why are hydrogen and oxygen studied before water?
   B. Why do you think the main divisions of plant and animal life are presented early in the book?

8. Criticism of the adequacy, correctness, or relevancy of a situation, statement, or diagram

   A. What was your main source of error?
   B. From your observation, what criticism should you offer of the posture diagram?

9. Inference from data

   A. From the facts presented, what do you think will be likely to happen to Victoria Falls during the next million years or so?
   B. What do you conclude concerning the relative rate of heat radiation from a dull and from a shiny object?

---

[1] adapted from Curtis, Francis D. Science Education, September, 1943.

10. Discussion

A. Present arguments for and against the introduction of the metric system in place of the English system in this country.
B. Discuss photosynthesis.

11. Outline

A. Outline a method of removing tarnish from copper.
B. List in order the various blood vessels through which the blood would flow in circulating from your toe to your heart and back.

12. Explanation or definition

A. What is meant by persistence of vision.
B. Tell in your own words what erosion is.

13. Simple recall

A. What is the chemical formula for water?
B. How many legs has an insect?

14. Summary

A. What characteristics do all satellites of the sun have in common?
B. What facts can you state about lines of force?

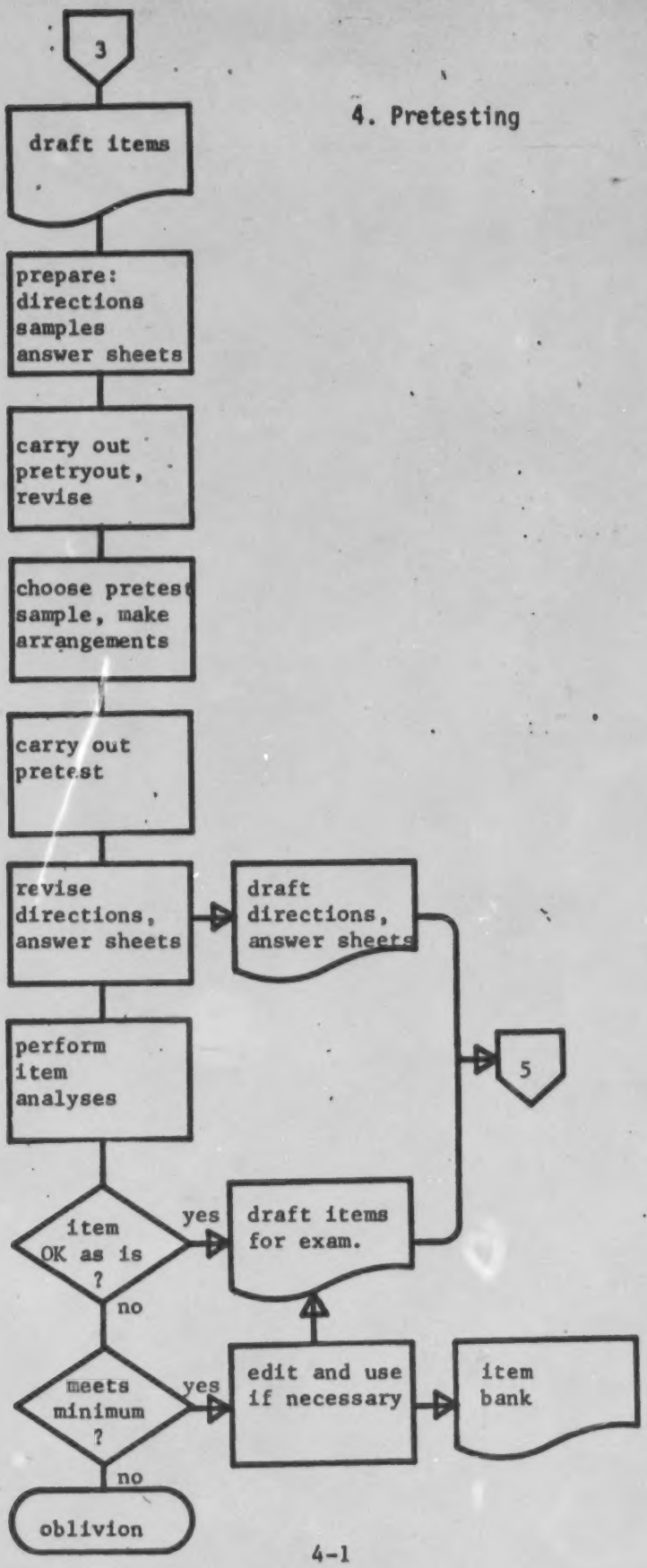
15. Observation

A. Where does the diagram show most of the planetoids to be?
B. What are the characteristics of the oxygen which has just been made by heating potassium chlorate and manganese dioxide?

16. Formulation of new questions

A. What questions occurred to you while performing this experiment?
B. Write three additional questions based on this experiment that you would like to have answered.

4. Pretesting



PRETESTING

Once the items and directions have been written, the answer sheets designed, and all have been edited by both subject matter and test specialists (and, ideally, examined by an editing panel), they are ready to go through the various tryout procedures which make up "pretesting." The importance of this stage in the preparation of the examination cannot be overemphasized; it is here that the most effective ways of preparing and conducting the real examination are discovered. In fact, the inclusion of this step often increases the reliability of the examination as much as from .60 to .90, and thus bears much of the responsibility for the higher reliability of objective--compared with traditional--tests.

The reasons that pretesting increases the validity and reliability of an examination are perhaps too obvious to discuss at great length. A brief summary, however, may be valuable.

1. Through pretesting, weak or defective items are identified and discarded: those which are ambiguous or indeterminate, that have non-functioning or implausible distractors, that are excessively easy or difficult, etc. These items are often identified in unusual ways. After one pretesting session, inspection of the scratch paper used in an arithmetic test disclosed that the methods candidates were using to solve the problems bore no resemblance to those in which the syllabus committee had hoped to measure skill. In this case, the items were revised so that the correct answer could be consistently arrived at *only* by use of acceptable methods.

2. Another argument for pretesting is that the difficulty of each item is assayed, and thus the overall level of difficulty of the entire examination can be more accurately estimated. In the traditional essay test, the difficulty of the items is problematical. Typically, the marks are adjusted after the scripts have been read so that the correct proportion of candidates pass. Using pretested objective items enables testers to set suitable difficulty levels in advance. (The term "difficulty level" is replaced by the more logical "facility level" in the U.K.)

3. The power of each item to discriminate between those who know the material and those who do not can be determined by pretesting. When the items are chosen which prove to be the most discriminating, the examination increases in efficiency per unit of time, and in general has higher validity.

4. Pretesting provides data needed to determine how many items are

necessary for the finished test, and the appropriate time limits. At least this is true in theory. In the reality of testing in many countries, the test committees or the ministries decide the number of items and the length of time to be allowed. However, the pretesting does provide data which help them prescribe limits.

5. In addition, weaknesses or needed improvements are discovered in the mechanics of the test administration: directions to the supervisor and candidate, provision for responses, sample exercises, typographical format, and so on. In one pretesting session, for example, candidates from one area would not turn the page unless specific instructions to do so were written at the bottom of each page. Some words had to be changed: "capital letters" to "block letters," "pitcher" to "jug," and "water glass" to "tumbler." Once a reading comprehension passage was printed on one side of a page, and the questions concerning it on the other. Candidates were turning pages furiously, wasting time and energy. On another occasion a particular color of print and size of type were found unsuitable in an area where eye infections were common; a similar discovery was made during testing in unlighted buildings during the rainy season.

In writing directions to supervisors, it is necessary to anticipate most of the questions they and the candidates will ask. Adherence to time limits, methods of marking answers, behavior for emergencies, handling of latecomers, and return of materials, should be discussed in detail. It is also wise to supply embarrassment-avoiding answers to sample questions. Ideally, of course, the candidate should be able to proceed correctly from either his own instruction booklet or the supervisor's instructions alone. That this is desirable, and sometimes even necessary, was demonstrated dramatically in Nigeria. Candidate information booklets had been sent to every single student about to take a secondary entrance examination. Supervisor directions had been carefully drawn up but not trial tested. When the examination date arrived, officers of the STO decided to make a spot check at some centers to see how well the examination was proceeding. At one center, the administrator was giving completely erroneous directions. He was not reading what was in the instructions, and did not understand even the most elementary principles of testing. Listening outside his room, the officer was horrified, and despaired of the results. Upon entering the room, however, he found that everyone was answering properly despite the

administrator. As the STO later found out, the candidates had all read the information booklet, and so were not misled.

As a result of this experience, the STO changed the supervisor instructions and tried to prevent a recurrence. But keeping each set of directions sufficient in itself is still the best safeguard. (A sample of supervisor directions is shown starting on page 7-9).

As important as these functions are, pretesting clearly does more than increase validity and reliability; it simplifies all aspects of test making and administration.

In addition to the regular directions which are being pretested, the STO should prepare a paper outlining the purposes of pretesting. This would be given to all those assisting in the examinations, and would include the need for security, the special time considerations, the need for comments on methods of improving the examination, and the common courtesy of thanking the principal or headmaster for allowing the trial testing at his school.

Some of the problems which pretesting identifies can be uncovered even earlier than the actual trial test. A procedure which is highly recommended for spotting gross deficiencies is the pretryout. This is a preliminary administration of tentative tryout units to small samples of examinees, after which, typically, no statistical analysis of individual items takes place. From three to ten subjects, similar to the testing population, are brought into the office. The items can be presented singly on cards, or they may be photocopied or duplicated in test format. A senior staff member administers the items, using the actual directions which are intended for use in the examination.

Through the pretryout phase, the emphasis is on spotting such things as major omissions, ambiguities, or inadequacies in the directions. Difficulties in the sample items are often noticed here. If the pretryout papers have been reproduced in a manner similar to that which will be used for the final printing, it is often possible to identify problems in such matters as the assembly of test booklets, or spaces on the answer sheets.

On occasion the STO staff have asked the candiates to think aloud while taking the test so that they could detect such things as difficult vocabulary or sentence construction, unexpected hints or hindrances to the right answer--in short, obtain unsophisticated reactions to their questions. They had already had expert opinion; this gave them the advantage of a view from the

other side. A candidate's on-the-spot reactions were followed up by an interview when he had finished.

Although these last two procedures have obvious merits, they have one disheartening side-effect: they are ego-deflating to the testers. Often staff members found that the candidates got the correct answer through erroneous reasoning, or at least *not* by the method they had in mind when they constructed the item. At other times the examinees cast aspersions upon some of their most cherished questions, ("This is a stupid one. I can choose the right answer without reading the question!"), and since everyone was so ego-involved, the temptation to argue was almost irresistable. Argument, needless to say, would have altered the subsequent responses, and caused the examiners to receive erroneous information.

In compiling the test before trial testing the following points need to be checked:

1. The title of the paper, number of questions included, and the time required for it should be clearly stated. (But see the caveat on page 4-11).

2. Examples given at the beginning of sections should involve the same process as the examinee is asked to carry out in that section.

3. The instructions and the alternative choices should be clear and unambiguous.

4. There should be one and only one correct answer, and the correct answers should be randomized; i.e., each answer position (A-E) should appear about the same number of times and none should appear more than three times consecutively.

5. Where alternative answers complete the stem, each alternative and the stem should form a correct sentence.

6. There should be no leading words in the question; that is, words which will give away the correct answer.

7. Answers which complete a stem should not start with capital letters unless of course they begin with a proper noun.

8. The lettering on any diagram should correspond with references made to the diagram in the questions. The letters A-E must not be used on diagrams since these letters are used for the alternatives.

9. Commas, question marks, and periods should appear in their appropriate places, and hyphenated words and those which must be spelled with a capital letter, such as Bunsen burner, Fallopian tube, should be correctly spelled.

Once the gross deficiencies are corrected, the next step is to locate a suitable large sample of candidates to trial test. The size and characteristics of the sample must be carefully considered. If errors are made at this point, no amount of statistical gymnastics can make a success of the item selection procedure. The best method by far is to enlist the help of the Ministry of Education. For a national educational examination, a collection of about six to nine schools from all parts of the country, made up of schools rural and urban, good and bad, segregated and coeducational, is generally required. The best results seem to come from having a minimum of 100 examinees answering each trial question.

In Nigeria the research section of the office checked the accuracy of the item statistics as estimates of the parameters.* To do this, they ran item analyses on a population of over 50,000 tests taken by secondary school applicants one year. Calculation of the actual standard errors showed that the sampling method was good; the σ for the difficulty indices was about .05 and for the discrimination indices .07. Repetitions of this method of checking produced similar results. They were able to achieve such accurate indices only because of the careful selection of the trial subjects.

In order to pretest a large number of items in the chosen sample, the pool of items must be divided into several similar groups or subsets. Each subset is then printed in a separate booklet, each booklet generally referred to as Form A, Form B, or Form C, etc. During the trial testing, two forms are used simultaneously in the same examination hall, having been distributed alternately to examinees. Directions about how this should be done must be given, preferably followed by a diagram such as this:

B A B A B A B A B A B A
A B A B A B A B A B A B
B A B A B A B A B A B A
A B A B A B A B A B A B

(distribute the papers so that alternate examinees receive different forms of the test.)

Each examinee is therefore answering only the items contained in one form. This arrangement helps to preserve the security of the final examination, since each examinee is exposed to only a fraction of the items during the

*The term "parameters" refers to the item analysis values in the total population, in the larger domain; the term "statistics" refers to the values in the sample, which is a subset of the population. If the sample is chosen properly, the statistics will be good estimates of the parameters; i.e., they will be within certain bounds "most" of the time.

trial test, and only a part of the whole examination is given within any one examination hall. For complex statistical reasons, to the extent that the forms are homogeneous, the methods of item selection are improved, so items should be placed in the different forms on the basis of similarity of subject matter content. Some offices include a small number of common items in all forms of the tryout tests. Statistics on these common items are compared for the different groups to determine if the groups are in fact similar.

Another difference in directions for the pretesting is having the candidates use the number on the test booklet as their student number. If each form of each paper has a distinct number, the answer sheets can easily be separated at the conclusion of pretesting. To illustrate:

Paper	Form	Numbering
A	A	100-199
	B	200-299
	C	300-399
B	A	400-499
	B	500-599
	C	600-699
C	A	700-799
	B	800-899
	C	900-999

The other difference between pretesting and the actual examination is the matter of time limits. Estimating the time required to administer the completed examination from data gleaned in the pretryout phase is a difficult matter. Time limits in the trial testing must be great enough to ensure an adequate tryout of all items. If many examinees fail to try an item because of lack of time, it is impossible to estimate accurately how many would have been able to answer correctly. But for administrative reasons, it may not be possible or practical to allow time for all examinees to finish.

Time limits are therefore suggested for each of the papers used for trial testing. The supervisor and invigilators should observe the examinees throughout the testing period. If all have completed a paper before the suggested time has expired, the signal to stop may be given, and the supervisor proceed to the next paper.

However, if examinees are still working when the suggested time has elapsed, the supervisor must decide, depending upon the circumstances, how much additional time to allow. On the one hand, he must keep in mind the

total time available for all the papers, for it could be uneconomical to sacrifice one of the later papers only to allow time for a minority of the examinees to finish. As a general rule, the supervisor should call time a few minutes after he has determined that about 90 per cent or more of the examinees have *attempted* all items. At this time it is imperative that he write down the length of time used on that particular paper, and that he does this for each paper. The STO can then take this information into account in setting the time limits for the actual examination.

The decision to call time on a paper depends on the judgement of the supervisor. No specific rules can be given except that he must make the best compromise to ensure that:

1. in the total time allowed, the maximum number of examinees is permitted to attempt the maximum number of items;

2. all papers are administered in such a way as to meet the objectives above;

3. the maximum motivation of the examinees is maintained throughout the testing period. If the motivation of the majority of examinees would be seriously impaired by their sitting idle for a long time waiting for only one or two others to finish, it would be better to call time and go to the next test. Obviously it will be of no value to force most pupils to wait for long periods if the result will be an uninspired performance on subsequent measures. Thus the examiner must be sensitive to the examinees' reactions and act accordingly.

This method will clearly not work on speeded tests, yet completion of almost all the items is necessary for the best results in item analysis. If a time-limited test is given without limits, results are often inconclusive since the mental set interacts with the rate of work of the examinees. One way out of this apparent impasse is to give the paper with strict time limits, but to include a set of "cushion" items of the same type at the end of the trial items. These are not marked, of course. A second method is to produce two forms of the paper with different orders of items. This, however, increases the expense and administrative problems of the pretesting, factors which dissuade many testing organizations.

The third method is to set a fairly generous time limit calculated to allow all but a small proportion to finish, and then in item analysis to make statistical corrections for guessing and non-attempts. If too few examinees

have tried a particular item, hoever, no amount of statistical manipulation will compensate.

Another technique is to intersperse new items in current examinations, as is done in many developed countries. This eliminates a great deal of expense, and would seem appropriate at first glance for all trial testing. There is only one drawback, but it is crucial: this method requires that every single test booklet be returned, uncopied. The new items could be inserted in a small subset of the whole population, and testing officers could be assigned only at these places; conceivedly this would preserve security.

As outlined earlier, the items are divided, assembled as different forms of each paper, and then reproduced. Since the particular combination of items exists only for the temporary purposes of pretesting, it is unnecessary to use anything more elaborate than mimeographing for the reproduction of the papers. (See section six on methods of reproduction.) Some offices are fortunate enough to have a method of producing inexpensive paper offset masters which provide clear copies from typed or handwritten copy.

A good rule of thumb for deciding how many items to include in a single form is to assume that a primary school leaver can answer one objective item per minute. Using this formula, an office may put about 50 or 60 items in each paper, and allow slightly more than an hour. Using fewer than 40 means that the total score used in item analysis may not be reliable enough to give an accurate indication of the worth of each item.

After the item writer makes an estimate of the difficulty of each item he submits, the items are placed in each form of a particular paper in ascending order of difficulty, both on the pretest and on the final version. The rationale for this is that such a practice will help the less secure candidates get started with a minimum of discouragement, and all the candidates answer a maximum of items. Some test makers rearrange items for *each* form to ensure that a sufficient number of examinees attempt each item, but the extra difficulties of keeping track of all the forms soon becomes overwhelming for multi-paper tests, and using the time limits mentioned above gives sufficient numbers for item analysis.

However, the wisest course is never to give all possible forms of any one paper at a single center. To do so would be to risk the loss of an entire subject paper if security is lax. This happened at one school in

Southern Africa where the methods of administration suggested above were not adhered to, and local teachers were used as both administrators and invigilators. To compound the error, the testing office only made two forms of a paper, and tried them both out at the school. The teachers took multiple copies of both forms and the testing office was left with no items in that subject.

A very common problem occurs in the collating of the various pages making up a form of the paper. Sometimes a page is blank, but the most serious problem is that occasionally an entire page is missing. For example, if page 5/6 is missing from a ten-page paper, and this is not spotted before the pretesting takes place, the testing office will never know if the page was actually missing or if an enterprising examinee tore the page out of the booklet.

A final note on the format of the trial test has to do with arrangement of correct answers on the sample items. No matter how much effort is expended on the mechanics of keeping the answer sheets of the many trial forms separate, a mixup always seems to occur. If the different forms are laid out so that the pattern of answers to the sample items is different on each form, another source of identification is available in questionable cases.

Motivation is not as much of a problem in developing countries as it is in others. Examinees of all types look upon the opportunity to attempt an examination similar to the one which might decide their future with more motivation than sometimes seems desirable. In spite of this, a common practice is to inform them at the beginning of the test that they should do their best and perform as if they were taking a final paper. Sometimes one can tell them that the results from the trial will be compared with their results from the final paper and inferences will be drawn. No one knows what this means any more than they do, but it has a most satisfactorily portentous ring.

Not only the examinees, but the school personnel are so motivated that security becomes a problem. Since attempts to retain copies of the trial tests by teachers and headmasters are even more common than by students, the importance of developing the best possible security measures is paramount.

The only way to maintain security is by absolute control of every single document. (Details of security controls when printing the examination are given in section six on reproduction.) One obvious method of doing this is to number all tests. After bitter experience with hand numbering, one office

changed to the use of an automatic stamper which advances one unit after every impression. The supply of papers is counted at least twice in the printing room, and then brought to a clerk for numbering. Since he is not told the number of each type of paper, his numbering the last one the same as the print room count provides the first evidence that all is well.

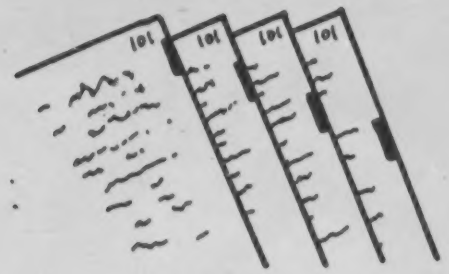
But accepting the count as correct would be a serious mistake. When ready to trial test, it is necessary to check the stamped numbering for accuracy; i.e., every paper must be numbered, and all in sequence. In more than 15 separate pretestings in the author's experience, no numbering has been found correct (and it was worse with hand numbering). Experience urges pessimism and caution.

This method should be carried one step further: when every page in every booklet is numbered, single pages cannot be taken. In one trial testing session, candidates were able to obtain copies of every paper by the simple expedient of organizing to get single pages. The first man took page 3 (4 on the reverse side), the second page 5, the next 7, and so on. The remaining candidates undertook the job of remembering the questions on page 2. Page 1 was directions, so they had it all.

In addition to numbering every page, it is most helpful to print a small stripe along the edge of each page, a stripe similar to a guillotining mark in printing. The first page has a stripe extending from the top right-hand corner of the page down one inch. The stripe on the next page begins where the first left off; that is, it begins one inch down the page and ends two inches from the top. This system extends throughout the paper. (See Figure 4.1 below.) As the tests are handled it is easy to see if all pages are printed and assembled correctly, because if all is well the stripe is solid. If a single page is missing, its absence is immediately evident.

Another way to ensure security is to refrain from advertising, a difficult task for the civil servant. The tendency is to label the first page of each paper something like "TRIAL TEST OF ARITHMETIC REASONING FOR THE 1965 SECONDARY SCHOOL ENTRANCE EXAMINATION," a clear call to arms. If possible, try to be overly modest and place some hieroglyphics such as "034/65/a" on the papers with no other identification. For the first few years after one office changed to this system, the candidates refused to believe that the papers could have anything to do with the ministries.

Figure 4.1 Edge stripes to prevent loss of single pages.



Of course just ordinary, determined vigilance is essential throughout the pretesting. If the supervisor distributes the papers in numerical order and personally collects them in the same order, he can promptly locate the probable culprit if a paper is missing. Additional directions for trial testing often say:

> At the conclusion of the testing time, you should collect the question papers and count them carefully. *Every one must be returned.* You should also note the beginning and ending times of the testing. If, when you have finished collecting and counting the papers, there are some missing, you should inform the examinees and refuse to proceed until *all* the papers are located.

Surprisingly enough, this last method always seems to work. A student, or more often the headmaster or an invigilator, will come forward with a paper shamefacedly, and proffer some weak excuse.

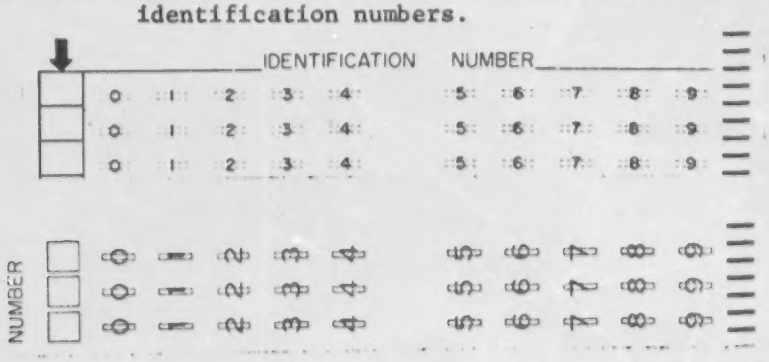
In some countries the STO may say, "Under no circumstances may anyone other than the pupils look at the question papers." In others, "If a member of the school staff wishes to see a copy, he should be shown one. He must not take it out of your presence nor must he have any scrap paper with him." Whenever school officials enter the room, extra attention is necessary, yet without their cooperation, pretesting would be impossible.

As is obvious, security depends upon the individuals administering the pretest. In many countries, the administrative arrangements for both the pretesting and the final examination are similar. The advantages in such cases are that the system is checked out well beforehand, and the whole procedure is much less expensive than when the senior or intermediate staff of the testing office have to be used. In spite of these considerations, however, preference is for the use of the testing office staff for the pretesting. Security is much tighter, and everyone from the testing office is much more attentive to ways of improving both the examination itself and

its administration. One compromise which could be worked out is that testing staff serve as invigilators, while local officials or teachers act as supervisors. In this way the staff members can pass among the candidates as they fill in their examination numbers, assist in answering their specific questions, and keep a close tally of the test materials. They can also note the frequency of procedural errors among the candidates, the number of candidates who need special assistance, and the amount of hesitation at the starting signal.

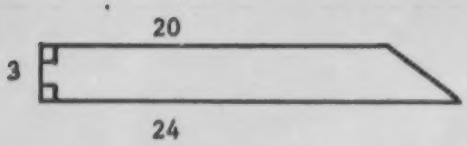
Another aspect of the entire examination process which is included in the pretesting is the adequacy of the design of the answer sheets. Every year some changes are made, sometimes major, in the design. At one time an office experimented with having the candidates grid their examination numbers in a horizontal direction, and when this produced too many errors, they tried having them turn the paper to a vertical position, which decreased the error rate (see Figure 4.2 below). The latter method is now being used on both answer sheets and answer cards (see Figure 6.3).

Figure 4.2 Horizontal and vertical formats for candidate identification numbers.



When all the pretesting data have been collected, all the papers of each form are separated from the others and put together for item analysis. The analysis is carried out so that items which prove too hard or too easy or are insufficiently discriminating between candidates who score high and those who score low on the tests as a whole are discarded or revised. In this way the final form of the test will consist of items which have been demonstrated to be appropriate in difficulty and discriminating power.

Consider, for example, the following question which was included in a larger collection and trial tested:



The perimeter of the figure shown above is

A. 47
B. 50
C. 52
D. 72
E. impossible to determine from the data given.

Papers of the 100 candidates who took the trial examination were separated into two piles: strong candidates and weak candidates. Then the answer choices were tallied for each group separately, giving the following results:

Candidates	Choices						
	A	B	C*	D	E	Omit	Total
Strong	1	0	30	2	15	2	50
Weak	5	0	6	3	24	12	50

*correct answer

The difficulty index (or facility index) or an item is the proportion or percentage of the tryout group that marked the item correctly; the larger this number, the easier the item. (In this example where C is correct, the index is 36/100 or .36.) An objective examination meant to rank candidates in terms of achievement or ability should provide a range of difficulty in its items, so adjusted to the capacity of the candidates that no candidate will make either the maximum or the minimum possible score. If ten candidates receive perfect scores on a test, it does not necessarily mean that they are equal in ability. It may mean, and in fact it often does mean, that the test does not contain a sufficient number of difficult items to differentiate among the ten candidates. If, on the other hand, these individuals receive the minimum possible scores on the same test, it does not mean that they are equally lacking in the skills being tested; it often means that there is, on the test, an insufficient number of easy items to differentiate among the candidates. It can therefore be said that at both extremes of difficulty, items cease to discriminate and an item which every candidate marks correctly, although demonstrating mastery, contributes nothing to the ranking of the candidates. Neither does one which everybody in ignorance answers incorrectly, although there are times when educators need to identify areas of mastery and ignorance.

For most tests it is desirable that a majority of candidates' scores

be scattered in the middle section of the score range and that the frequency distribution of the scores on the test, when plotted, gives a normal or near-normal curve; average ability is normally expected of most candidates. But for one reason or another, a test may be designed within specified limits to produce a skewed, rather than a normal distribution. For example, in a secondary school entrance test where the purpose is to identify and rank the top ten per cent of the candidates, a very difficult test is desirable and the concentration of scores in such a test will be low on the score scale. In such a case, the poorest candidates may not be adequately differentiated from the average candidates, but the very best candidates will be clearly separated and properly ranked. With this in view, the average difficulty index of a test should be similar to the ratio of the number of required candidates to the total number of candidates. For example, the average difficulty index of a test for a large number of candidates in which one-quarter of them is required should be about .25.

A second important quality to be considered in selecting items for ranking candidates accurately in terms of achievement or ability is the discriminating power of the item. This is measured by the extent to which it separates those high on the criterion measure from those who are low. In the example on page 4-14 a simple index of discriminating power is the difference in proportions between those marking the answer correctly in the two groups; i.e., 30/50 - 6/50 = .60 - .12 = .48. The higher this value, the better the item is discriminating between the strong and the weak candidates. An item which both groups answer correctly in equal proportions contributes nothing to the discriminating power of the test. One which candidates with low ability answer correctly in a greater proportion than those with high ability contributes to an improper ranking and therefore has a negative discriminating power. An item may function in these negative ways for various reasons. It may be ambiguously worded, or it may have several defensible answers although only one is keyed as correct. But whatever the reason for the lack of positive discriminating power, such items should be stricken from the tests.

An example of a poor item which shows negative discrimination for the keyed choice D is:

> There are nine doves perched close to each other on a fencewire. A boy throws a stone at them and kills one. How many doves are left?

A. 0
B. 6
C. 7
D. 8 (keyed correct)
E. None of these.

Many reasons could be postulated for the tendency of the high ability examinees to choose other than the keyed answer: the word "left" is ambiguous since it could mean "left on the fencewire" or"left alive;" several answers are defensible; etc.

In assessing the discriminating power of an item, the criterion used to judge its value often constitutes a problem to test makers. Ideally, an external criterion or an independent measure of the skill to be assessed should be available but in academic achievement testing, this is very difficult to find. In practice, therefore, one usually resorts to an internal criterion: the total score on the test of which the item is a part. This practice is based on the assumption that a well-planned and well-constructed test does, as a whole, measure the skills which have been selected and defined by its maker and that these skills are the significant ones. In this case, a performance on the total test is taken as a proper measure of achievement and the test itself is taken as the criterion against which the discriminating power of individual items is judged. This may not be very satisfactory in theory, but is often necessary and justifiable in practice, and most methods of assessing the discriminating power of an item are based on it.

A number of methods of determining the difficulty and discrimination indices are available, and the choice of methods usually depends on the availability of calculating equipment.

The simplest of these are approximate methods, easy and accurate enough for amateur test makers with little knowledge of statistics. They are all based on a principle which can be illustrated as follows: Suppose a large number of candidates mark the answers to each of the 100 items in a trial test. The supervisor then scores the test and arranges the answer sheets in order of merit by size of the total scores. He then picks from the pile of answer sheets the top quarter and bottom quarter of the scripts and marks on another sheet the number of correct answers to each item in the two quarters. When this has been done for all the items, he goes through them one by one and discards those items with approximately the same number of tally marks in the top and bottom groups or those with more tally marks in the bottom quarter than in the top quarter. This process constitutes a simple method of item

analysis.

More complex methods of deriving discrimination indices are usually preferred since they have better statistical properties. Many of these are treated in textbooks on item analysis and they include the determination of the biserial, the point-biserial, the tetrachoric, the phi, and others. The normal procedures for obtaining these correlation coefficients require a great deal of laborious calculation, although ABACS are available. However, they are regarded as the standard procedures in item analysis, and the adequacy of other methods for determining discrimination coefficients is usually judged by the degree to which they are capable of yielding results which approximate those obtained by using one of these methods. Calculating equipment is now being used for obtaining these indices, so that one seldom has to go through the tedious calculations by hand.

The upper and lower group method is a quick method for estimating the discrimination index of an item when the other methods cannot be used for some reason. It is very popular with many test makers and was the method most often used before the advent of the computer. It requires the setting up of two groups on the basis of the total marks scored by each candidate in the test. The first group will be a certain percentage of the candidates (e.g., 27%) scoring the highest marks in the test, and the second group will be the same percentage of the candidates scoring the lowest. These, for obvious reasons, are referred to as the upper and lower groups.

In determining the discrimination of an item by this method, the number of candidates answering the item correctly in each of the two groups is counted, and these numbers are used to obtain the coefficient of the item from specially prepared tables. There are a few modifications of this procedure and different versions of the tables are in use, but two commonly used are:

John Flanagan, Calculating Correlation Coefficients. Pittsburgh, Pennsylvania: American Institutes for Research.

Chung Teh Fan, Item Analysis Table. Princeton, New Jersey: Educational Testing Service.

These tables can be used to determine both the difficulty and the discrimination indices of an item.

The only feasable alternative to the use of these tables is to carry out the item analysis by computer. A variety of item analysis programs are available, often free, from data centers in the U.S., the U.K., and

Australia. For STO use, a special item analysis program was written, and the output set up so that the computer printed the results on adhesive labels. The labels were then attached to the item cards for use in editing, moderation, and assembly into the final form of the test, or for placement in the item bank.

Since most test experts agree that the point-biserial correlation is the most appropriate statistic for assessing the discriminating power of an item, it was used in writing the computer program. The program consisted of a set of detailed instructions to be followed by the computer in carrying out the necessary calculations for the item analysis. The sequence was that:

1. Each of the answer sheets was scored and the responses of every candidate, plus his total score, were punched into a Hollerith card (or w written on magnetic tape).

2. When all the cards were ready, the program, and card containing general information about the data, and the data cards were all fed into the computer. The computer first read the instructions, then the data, carried out the necessary calculations, and then printed out the results.

The results consist of the mean score for all candidates, the standard devation, the reliability coefficient, the number of scores on which the analysis is based, and statistics on each item and alternative in the test. The actual form of the output is shown in Figure 4.1.

The first two rows show the examination title and basic statistics for the total score. Subsequent rows give details about the candidates' responses to each item. The top row for each question includes the difficulty index for each alternative; the second includes the discrimination index for each alternative. In item 1, for example, the difficulty of choice A is .426 and the discrimination -.357; the corresponding figures for alternative E are .065 and -.189. The last pair of numbers in each row shows the figures for the candidates who did not attempt the item, either because they did not know the answer, because they had insufficient time, or because they chose to omit it for other reasons.

Following the figures for each item are one or more explanatory sentences. These are based on the following general guidelines used for improving the examination:

1. The discriminating power of the correct answer for the item must be positive and should not be less than about .20. If this condition is not met, the item is generally discarded.

Figure 4.3 Sample item analysis data

```
ABC TRIAL TEST FORM B     OCTOBER 1970
NUMBER OF CASES = 261     MEAN = 29.1374     SD = 9.2931     R(TT) = 0.821

ITEM 1        A         B         C*        D         E        OMIT
  DIFF      .426      .065      .344      .000      .065      .100
  DISC     -.357     -.096      .678      .000     -.189     -.039
***RE-EDIT - ONE OR MORE ALTERNATIVES NOT FUNCTIONING

ITEM 2        A         B*        C         D         E        OMIT
  DIFF      .180      .114      .032      .213      .049      .412
  DISC      .270      .586     -.037     -.064     -.023      .020
****REJECT - TWO PLAUSIBLE ANSWERS
***RE-EDIT - ONE OR MORE ALTERNATIVES NOT FUNCTIONING
**RECONSIDER - ITEM EITHER TOO EASY OR TOO DIFFICULT

ITEM 3        A*        B         C         D         E        OMIT
  DIFF      .258      .112      .161      .048      .306      .115
  DISC      .468     -.067     -.027     -.011     -.261     -.213
ACCEPTABLE

ITEM 4        A         B         C         D*        E        OMIT
  DIFF      .023      .035      .023      .845      .063      .051
  DISC     -.181     -.161     -.166      .602     -.204     -.050
**RECONSIDER - ITEM EITHER TOO EASY OR TOO DIFFICULT

ITEM 5        A         B         C         D*        E        OMIT
  DIFF      .612      .122      .020      .142      .040      .084
  DISC     -.121     -.181     -.005      .616     -.110     -.013
**RECONSIDER - ITEM EITHER TOO EASY OR TOO DIFFICULT

ITEM 6 . . .
```

2. The difficulty of the correct answer should be roughly between .20 and .80 in most cases (but see exceptions on page 4-15). If the difficulty index falls outside these limits, it may be possible to rewrite one or more alternatives and retest the item again at a later date. The proportion in which these items are present in a paper, and the average difficulty index of all the items, will depend on the purpose of the test and the proportion of the candidates to be selected for treatment.

3. The difficulty index of each alternative answer should not be less than .05. A value less than .05 indicates that the alternative is not functioning well and should be changed. However, if, as sometimes happens, the correct answer has a high discrimination index (e.g., over .4), this step may be ignored. Some authorities suggest that question writers always attempt to set five alternatives for each item. If after item analysis one alternative is shown to be not functioning, it can be eliminated and only four choices used in the final form.

4. All the distractors and the omits category must have a negative or zero discrimination index. A positive value indicates that the alternative is attracting upper group candidates and should therefore be replaced.

5. The number of candidates omitting an item must not be too large. Since this aspect is interrelated with the other conditions, it is not usually stressed.

Using the rules suggested above, the statistics in Figure 4.1 can be considered to see how many items can be regarded as suitable.

1. The correct answer to item one, marked with an asterisk, is C. If the selection rules are applied, one can see that choice D, with difficulty .000, is not functioning. Since no other rejection rule is applicable, and the discrimination of the correct choice is so high, one can either leave the item as it is or rewrite alternative D. The decision rests largely on the willingness of the staff members to try to improve an item which is already fairly sound statistically.

If the decision is made to rewrite alternative D, the writer might first inspect choice A. Since this alternative is working so well, the new D might be structured along the same lines.

2. Item two cannot be regarded as acceptable because:

   A. with a difficulty index of .114 for the correct answer, the item is apparently too difficult for the group;

B. alternative A has a positive discriminating index even though it is not the correct answer. The implication of this is that it attracted some of the upper group candidates and is therefore a bad distractor;

C. the proportion of omits (.412) is very high, indicating that many candidates probably could not choose between A and B and therefore left the item unanswered.

Some testing offices would rewrite alternative A and then try the item the next year in the hope that such revision would make it usable. Obviously such a procedure saves a considerable amount of effort.

3. The correct answer to item three is A, and a glance at the figures for this item will show that it is statistically acceptable as it is.

4. The most serious criticism of item four is that it is too easy. This also accounts for the fact that the distractors attracted very few candidates. A few items of this nature probably could be used as the first ones in an examination paper since candidates seem to do better in the entire paper when they gain confidence from answering relatively easy ones at the start.

5. The correct answer here is choice D, with a difficulty index of .142 and a discrimination index of .616. It is a very difficult item with a high discriminating power, and might be used toward the end of a paper.

After the best items have been selected in the manner shown above, the remaining non-rejected items are placed in the item bank for future consideration. At a later date, when another trial test is to be carried out for a similar type of examination, some of the items are modified and pretested again.

The methods of selecting items just outlined work for other purposes as well. In one country the STO was accused of discrimination in the secondary school entrance examinations. On the traditional examinations, candidates from one part of the country did significantly worse than those from another. Those who did badly came from an area which was populated by nomadic tribes, and which generally contained the poorest primary schools in the country. In contrast, the more fortunate students were resident in agrarian areas where the schools had been instituted centuries ago by missionaries. Because of the great variation in educational opportunity and quality, very few candidates were selected from the former, and a great many from the latter.

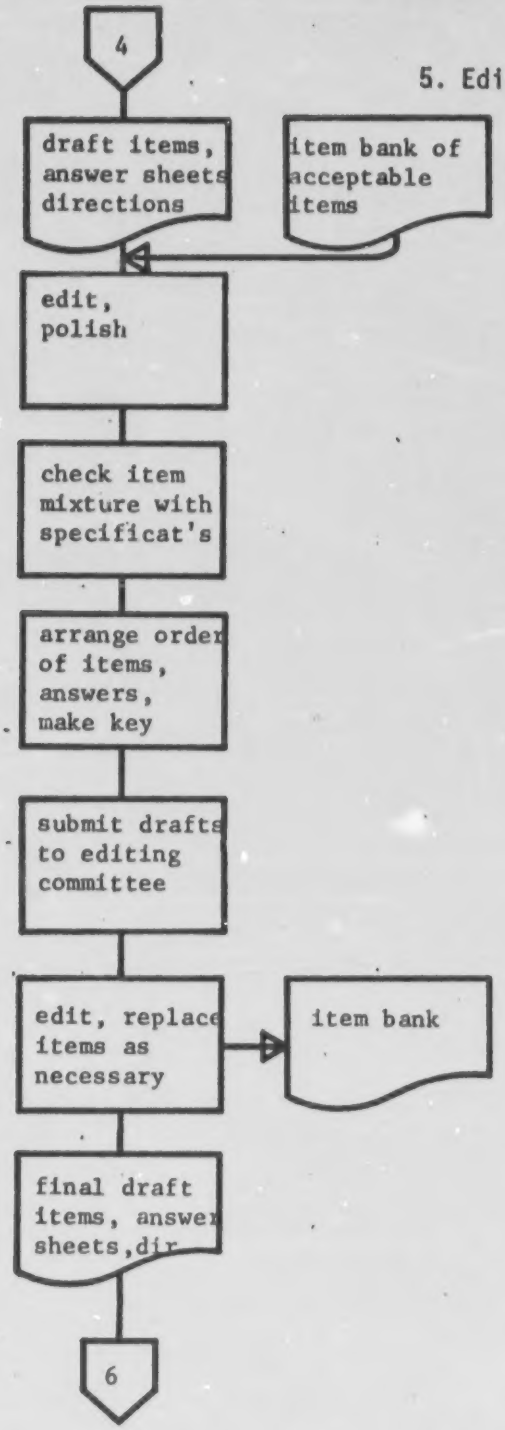
In order to correct this imbalance, the office added some aptitude tests

to the selection battery, attempting to uncover abilities less dependent upon the quality of primary school experience. They proceeded on the assumption that verbal and numerical aptitude were distributed evenly throughout all sections of the nation, so if an item showed roughly the same difficulty index for a sample of students from one area as for a sample from the other, and roughly the same discrimination index, it was considered fair and usable. If the item characteristics were too different, the item was judged as overly contaminated by education, and thrown out.

When the resulting examination was given, the proportion of candidates selected from the different parts of the country were approximately the same, and the charges of unfair discrimination were withdrawn.

In summary, the pretesting of questions provides a means of judging how well they will actually "behave" when answered by various candidates. The final form of the test then consists of items which have demonstrated that they are appropriate in difficulty and in discriminating power. A copious file of items of various sorts, with known characteristics for certain populations, soon accumulates so that new tests and forms can be made to order.

5. Editing


4
draft items, answer sheets directions
item bank of acceptable items
edit, polish
check item mixture with specificat's
arrange order of items, answers, make key
submit drafts to editing committee
edit, replace items as necessary
item bank
final draft items, answer sheets, dir
6

EDITING

Good items are laboriously made, not born spontaneously. They are written and rewritten at each stage of the cycle, from item preparation up to the reproduction of the final copies. Editing following the pretesting phase, however, is not as comprehensive as at earlier stages, since there is danger that any extensive changes may change the item characteristics. Such changes could then be detected only by pretesting and analyzing the items again. Unless the best items at this point are allowed to stand much as they are, the pretesting process could be repeated endlessly. So if the revision of an item endangers its point, it is either discarded or used without changes.

Following pretesting there are, therefore, hopefully no drastic modifications. Instead, this is the time in the examination cycle for the final polishing and arranging of the items before publication. If each individual item, its source, author, category, and record of changes in format, working, etc., are written on a separate notecard--as recommended earlier--the history of that item from its initial production to its final printing can be seen at a glance. Such readily accessible information helps the test maker decide the most effective order in which he might put the items in the test. For example, if a number of questions are taken from a particular reference source, or are based on the same table of logarithms, he might place them together in the examination. Sometimes the temptation seizes him at the last moment to modify an item, which, if he succumbs without consulting the history of that item, might cause him to revert to one of its earlier forms which was discarded for some inadequacy. Simply as a time-saving device, the notecard method is thoroughly justified; having data on notecards allows a simple reshuffling, and if necessary, endless rearrangement of items until the desired order is obtained.

The notecard should give not only the source of material from which the question was taken, but also the proof or authority for the right answer. Justification for including the author's name is that he can be identified and relied upon as a good source of such items thereafter. The category of the question--such as objective, essay, or behavioral objective 6a, 7, or level 1, level 2, etc.--is included on the item card so that all similar types of items can be located quickly to be placed together in the examination. The item analysis data also appear on this card. If adhesive labels are

used for the item analysis output, these are attached to the card; otherwise, these data are transferred to the card by hand. An example of a completed item card is shown in Figure 5.1.

When all information is on the cards, they can be sorted into order for the editing committee or for final reproduction. Assuming that a sufficient number of items have survived the item analysis process, there are several methods by which a final assortment of the best of these can be chosen and arranged. If the test is not tied directly to an outline or set of curriculum objectives, plotting the items using the item analysis data gives the best results (see Figure 5.2).

In this diagram, those items falling in the shaded area are probably unusable; those above the dotted line may be usable with some revision. The usual approach is to place the items in order of difficulty, the easiest first--in this case 16, followed by 8 and 9, etc. But a slightly better test will result if the easiest of the most discriminating items are placed first: 9, followed by 6, 15, 16, etc. In this way, the items are arranged to encourage as many candidates as possible to answer the most discriminating items.

This method is most applicable to aptitude tests, but unfortunately is not helpful with the typical achievement test. With achievement tests, it is necessary to go back to the test outline or list of objectives and choose the best items within each classification. If however, as sometimes happens, not as many items have survived item analysis as are required within a certain classification, the tester is in a dilemma. In the absence of an item bank, he asks himself if it is better to write and use new items without pretesting them, or to try to change drastically some of the rejected questions. Unfortunately, either alternative is unsatisfactory; hopefully he will start with more items next time. Experience has shown that on the average about 20 per cent of the original items survive conscientious review and item analysis, and woe to him who thinks he can consistently do better.

Although many offices choose to arrange the items within the paper by subject matter, other possibilities obviously exist. Local ministries often consider the subject matter method desirable because they assume that candidates will have to do less mental "gear shifting." One objection to it, however, is that an unwise candidate might spend too much time on a particularly difficult question in one subject matter area and thus lose

Figure 5.1 Completed item card.

1 2 3 4 5 6 7 8 X 10

grid cell: A3
used: '70 SSST
date: OCT.

author: L. Borgia

Laugh is to happy as cry is to ____

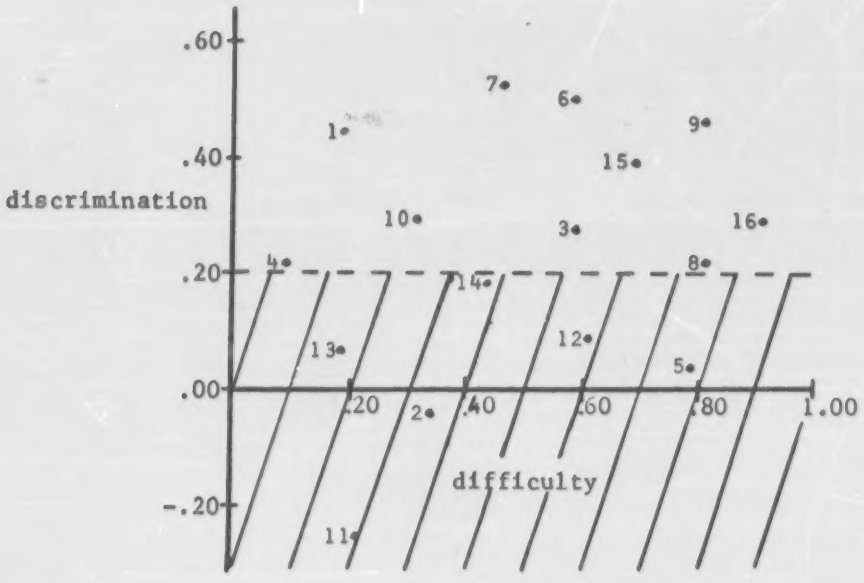
A. sad. B. wonder. C. hide. ~~D. rough.~~ E. lost.

D. baby (sj, 5/70)

n= 139
group: '69 ST m
keyed: A

	A*	B	C	D	E	OMIT
DIFF	.36	.21	.10	.00	.29	.04
DISC	.32	-.15	-.08	.00	-.16	-.05

Figure 5.2 Plot of item analysis statistics.



out on all the easy items in another.

Those who consider this a serious objection would probably be more comfortable if items were arranged by difficulty. Arrangement in this way leads to less confusion and discouragement on the part of the examinees and ensures that the maximum number of candidates answer each item. This, in turn, increases the reliability.

Another way items may be arranged is by type of question, a method which is naturally more applicable when a variety of types is used. In this format, all the true-false questions are presented together, followed by the multiple choice, and finally the matching. Some examinations which include several different types of objective questions could be arranged by a similar scheme; i.e., all verbal analogies, followed by reading comprehension questions, etc. This latter method becomes similar to arrangement by subject matter.

Another method depends on the classification scheme used for the items. If Bloom's taxonomy were used, all the level 1 questions would be first, followed by the level 2, etc.

No matter what method is used, care must be taken to distribute the correct answer choices randomly (thus evenly) among all the alternatives. If the test has four alternatives for each question, approximately the same number of A's, B's, C's and D's should be keyed as correct.

In situations where candidates' copying poses a problem, an STO may produce two separate orderings of the items and print two versions of the test. The items are the same, but printed in a different order. The two versions are used in the testing halls where pupils must sit two to a desk, and effectively discourages copying.

On the top of the test booklet, many offices indicate the time limits and number of questions to be answered. This helps the candidates plan their time more efficiently and is therefore recommended.

A final caveat to the compilers of items for final examinations: avoid overlapping. A candidate should not be able to answer one question from information given in another; nor should he be prevented from answering a question correctly by having failed to give a correct answer to a preceding one. If the stem of the question is long, it is more likely to give information which will help the candidate answer a subsequent one.

For instance:

31. Coulomb's Law of Electrostatic Attraction states: "The force of attraction or repulsion between two charged bodies is directly proportional to the product of the charges, and inversely proportional to the square of the distance between them." If F is force, Q and Q' are charges, D is the dielectric constant and d is distance, a mathematical statement of the law is . . . .

53. The relationship $F = \frac{QQ'}{Dd^2}$ which refers to charged bodies, was first suggested by the scientist . . . .

In this example, number 53 answers number 31, and number 31 answers number 53.

Another form of overlapping is that of using the results of the first question for solving subsequent ones. For example:

13. The conversation in the poem "The Telephone Conversation" probably takes place in

A. West Africa.
B. England.
C. some other part of Africa.
D. America.

14. What words in the poem give you this impression? . . .

If the candidate fails to answer question 13 correctly, he will probably also fail to answer question 14 correctly. However. there is far from a unanimity of opinion on the relative merits of the interlocking question.

Although the layout is important, the content is unquestionably more critical. In some countries, a review committee, sometimes called an editing, syllabus, or testing comittee, is given the responsibility of bestowing the imprimatur on the final product. They review the items for accuracy and appropriateness of subject matter content, technical merits, and judge the general overall quality of the examination.

The composition of the review committee has a direct bearing on the popular image of the test, and on the reputation of the testing office. If the members come from positions of authority throughout the country, they not only ensure high standards for the tests, but also act as public relations personnel for the testing office. When an office does an impressive job of presenting items to the committee, they advertise the high quality of the examination and the expertise of the testing office staff throughout the country more effectively than any newspaper advertisement ever could. Conversely, when they do a less than adequate job, they are publicly humiliated. The criticism in the latter case, although hard on the ego, is nevertheless valuable.

Obviously the review committee members need to understand the rudiments of item analysis, for when they do, they can appreciate both the effort which has gone into the pretesting, and the information which the item statistics makes available to them. Consequently, a five-minute explanation of the process--similar to the example presented earlier--can be given to any new members before consideration of the actual papers.

The presentation of the items and statistics can be made either together or separately. When staff members show them together, they usually use item cards. The committee looks over each item separately and then considers the composition of the paper as a whole. If they object to any item, the staff give them alternatives which they consider and can accept in lieu of it. Instead of the original item bank cards, they usually use some simpler ones for this purpose, containing only the item, the category, and the statistics; for political reasons they do not identify the author or the particular group used for pretesting.

On rare occasions the complete draft may be presented for acceptance, accompanied by duplicated tables of the item statistics for the committee's perusal. If the committee agrees to change any item, it is necessary to modify the paper by pasting in the new question (which, hopefully, is the same size as the former one.)

The former method of presentation to the committee is preferred to the latter method for a number of reasons:

1. The card method gives less of an impression of finality, and shows a greater willingness to respect the wishes of the committee.

2. It is physically easier for the members to consider an item when the statistics are presentedin direct proximity.

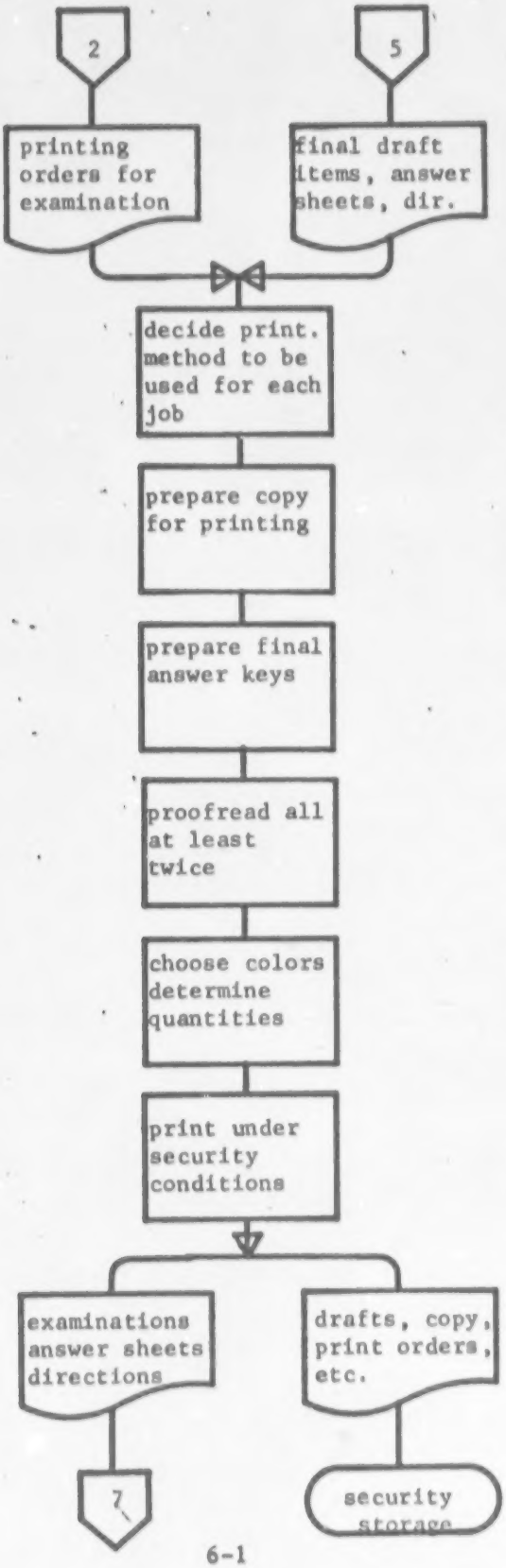
3. Presentation on cards may make it more difficult for committee members to remember the items, and thus preserve the security better. No testing office should underestimate the pressure sometimes brought to bear on committee members to discuss the content of an impending examination with relatives and friends. It is often phenomenal.

In regard to this last point, some offices also ask each moderating committee member to sign a security statement similar to that shown on the bottom of page 3-18.

Surprisingly enough, moderating the examination papers is a pleasant experience. Perhaps the fact that, until this time, most discussion has

been within the organization accounts for the stimulation and excitement often felt in discussion with the committee. Or perhaps the knowledge that this meeting represents the culmination of months of hard work in preparing and editing the questions, leads to rejoicing that the examination cycle is half over.

6. Reproduction



6-1

REPRODUCTION

A difficult compromise that has to be made in test reproduction is between beautiful appearance and low cost. Thousands of candidates will see the test booklets, and they should therefore be a good advertisement for the STO. However, the tests will generally be used but once and then destroyed. Somehow a compromise must be made so that as good an impression as possible can be made at the lowest cost.

Although an STO may have some equipment, it will rarely have what is needed for producing the vast quantities of tests needed in large-scale examinations. And the testing office has to reproduce more than just the final question papers. The answer sheets, directions, and accessory materials must also be duplicated, and at some time or another in the course of the examination cycle, multiple copies of the questions papers are required for editing, pretesting, and moderating. Since so many different materials must be reproduced for so many different purposes, the STO will probably find it practical to use different methods of reproduction depending on the quantity, desired quality, and intended use of the materials.

At present, four reproduction methods are commonly in use: photocopying, office-type duplicating, offset printing (lithograph), and letterpress printing. A general comparison of these four methods is shown in Figure 6.1. The time required to prepare the copy is almost the same for each mode of reproduction.

The photocopying technique is the reproduction of single sheets on an office-style copier. With most copiers of this type, one page is inserted and copied at a time. Each page takes from a few seconds to a few minutes to copy, and may cost between one to fifteen cents each. This method is ideal for the small quantities used in editing and moderating.

For somewhat larger quantities, the office-type duplicating machine may be adequate. In past years, the copy had to be typed on special stencil masters, but more recently developed methods allow the user to make a master directly from an ordinarily typed copy. In the latter case, the diagrams and pictures can be reproduced easily (except for halftones--shades of gray); in the former, such illustratory material would have to be copied by hand onto the master. The cost of producing copies with such equipment is the lowest of all the methods mentioned here, but this fact is offset by the less happy one that the appearance is the worst of all.

Figure 6.1 A comparison of four common methods of reproduction

	Photo-copying	Office-type Duplicating	Offset (Litho)	Letterpress
Typical Quantities	1 - 20	10-- 500	100+	1000+
Appearance	Good	Fair	Good	Excellent
Printing Time*	1 day	1 day	7 - 10 days	2 - 6 weeks
Size of Copy	Same as original	Same as original	Variable	Variable
Cost	High	Low	Less for small quantities	Less for large quantities

*approximate length of time needed to produce the typical quantities shown in the first line of an eight-page booklet.

The mimeograph is often used for some pretesting reproduction, but most offices prefer offset printing for the actual test because of its low cost, ease of set-up, and the comparatively short interval between draft and final copies. In addition, a 7 x 10 inch offset page may contain as much copy as an 8½ x 14 inch mimeograph one without loss of reading quality.

The printer sets up all the copy in type for letterpress printing. Diagrams, pictures, drawings, etc., are transferred to metallic or plastic blocks to be set in frames with the text. One of its great advantages over the other methods is that it produces an extremely sharp image. Some offices take advantage of this characteristic to save pages by specifying smaller type. However, since a comparatively long time is needed for setting up the type, the test is out of the STO for a disquietingly long period, which greatly increases the chances of a security leak. Some STO's with large budgets send their materials out of the country for printing in the hope that security will be strengthened. What actually happens in these cases is that the tests remain even longer away from their supervision, the length of lead time is increased, and the effects on security are dubious at best. For these and other reasons, offset reproduction is generally preferred.

In the offset method, the prepared materials, termed camera-ready copy, contain everything which is to be reproduced. Diagrams and other non-textual material may be drawn on tracing paper and attached to the sheet containing the text. Photographs and diagrams taken directly from

textbooks can be used exactly as they are found, or can be changed in size. Various preprinted dry-transfer type is available for use in setting up the material. These are extremely simple to use, come in a large array of fonts and sizes, and are relatively inexpensive. Some companies produce patterns useful for shading on charts and drawings, and edging for certificates; see Figure 8.3 for example.

Calculating the size of the camera-ready copy is difficult, so it is well to work out the size of the *finished pages* with a printer. He will then be able to specify the size of the prepared copy--the printed material, that is, since the margins are unimportant. Some offices use specially prepared layout sheets of white bond paper of 24 pound weight (about 110 gms). Pale blue guide lines around the margins, center marks and line counts simplify the typist's job considerably, but do not show on the final copy.

The camera cannot reproduce detail which does not exist, so the typing of the material to be reproduced must be even and dark. The best inexpensive method entails the use of an electric typewriter with a carbon ribbon. With a manual typewriter and silk ribbon, care should be taken not to change a ribbon in the middle of a page, and of course the keys should be cleaned before proceeding. Mistakes can be easily covered with a white opaque liquid, obtainable from art stores, which can be typed over after it dries. Larger areas can be changed by typing or drawing the corrected material on patches which are then glued over the erroneous material. If rubber cement is used for all gluing, the patches can be removed and attached again if the positioning is not quite right the first time. In addition, rubber cement does not cause wrinkling and excess cement can be rubbed off when it is dry.

In both offset and letterpress printing; the general practice is to print at least four pages at a time, which means that the final number of pages is a multiple of four. If the number of copy pages is not a multiple of four, it must be condensed or expanded, or blank pages must be inserted. The major considerations are, of course, legibility and economy. With offset printing, the various pages can be cut and pasted in different configurations until a pleasing and legible result is obtained.

Pagination takes a bit of forethought. If only a single page of directions is needed, the printing of the test proper can begin on the next page. In this case, a candidate would read the directions and wait for the signal to begin before turning the page. But if more than one page is

needed for directions and sample items, the test would begin on page four to prevent candidates from starting to work before the signal is given. In other words, it is advisable to lay out the examination so that candidates must turn a page before they begin. With untimed examinations, however, the test can start on any page since no error is introduced by a candidate's beginning too soon.

Although the page numbers can be printed at the top or bottom, centered or at one side, one method does have an advantage. If the test form and page number are printed in *large* type at the upper outside edge, the supervisor and invigilators will be able to check more easily that the candidate is working on the proper test. Anyone who uses this method needs to remember that the upper outside edge is on the left of even-numbered pages and on the right of odd-numbered pages.

Once the number of pages has been determined, the layout on each page can be modified as desired. The amount of white space (unprinted area) depends mostly on the length of the alternatives for each item. If the alternatives are generally short, two columns of questions may be printed per page. If they are long, printing in two columns would make the questions short and choppy. In some cases the order of the questions may be changed so that short questions are kept together. Certainly a diagram or paragraph should be followed immediately by the questions pertaining to it; the candidate should never have to turn a page back and forth to answer questions and to refer to the reference material.

The combination of two examination papers in a single booklet, when possible, results in ease of administration and economy in printing. In order to discourage the candidate from working on the wrong examination, one office printed the second paper upside down. That is, the second examination began on the back of the booklet and proceeded toward the front. Thus a candidate would have to turn his booklet upside down to work on the other paper, an action more noticeable to a supervisor than the mere turning of an additional page. In addition, the first paper was printed in red ink and the second in green, and these colors matched those of their answer sheets; the answers to the questions on the red examination were marked on the red answer sheet, the green on the green.

If separate answer sheets are not to be used, the layout of the test booklet must be modified. For ease in marking the papers, answers should be

arranged along the right margin and the candidate instructed to underline, circle, or check the proper answer. However, the system of marking in the test booklets would seldom be used for large-scale examinations as defined in this handbook.

Since the printing is usually done on both sides of a page, the paper must be opaque enough that the printing does not show through and confuse the candidate. Surprisingly enough, two types of paper show the best characteristics--the least expensive and the most expensive. The very high-quality cartridge paper is excellent in this regard, as is the paper used for newsprint. For one large examination where 200,000 sixteen-page booklets were printed, the difference in cost was almost two thousand dollars.

Printing of the administrative directions on an offset press would likely be too expensive an undertaking unless more than 500 copies were needed. For fewer than 500, duplication should probably be done on an office-type duplicator unless the office is confident enough of their directions to print up a five-years' supply--a highly unlikely eventuality.

Although newsprint could easily be used for the test booklets, it would prove unsuitable for the answer sheets. Whether they are to be marked by hand or by machine, they must be returned to the office clean and unwrinkled, and the paper which is most likely to return in this condition is a high-quality cartridge type. Any possible savings on paper would be completely offset by added expenses in scoring.

Some testing offices use the foolscap size paper for answer sheets, and print their own. Others use the 8½ by 11 inch size, especially if they use an optical mark scoring reader such as the IBM 1230, Digitek 100, or ICL UDT . In spite of the equipment manufacturer's statements to the contrary, it was possible to print the latter types of answer sheets in Nigeria. To do so, the copy was laid out with extra care, guillotine cut marks were placed at every edge, the timing marks were printed in the reflective color--i.e., non-black--and overprinted in black to check the registration, and the paper was allowed to absorb moisture before printing.

The use of answer cards, the size of ordinary Hollerith cards, has been showing good results in some places. The cards are more resistant to the rough handling of ordinary test administration, and can be fitted more easily to the writing surfaces in some of the more remote schools, which

in some places have been the ends of soft drink cases or small lap-boards.

The costs of the various types of answer sheets do not vary much when large quantities are printed. Typically the foolscap, 8½ by 11, or answer card cost will range from 0.3 to 0.5 cents each.

The design of answer sheets for machine scoring is discussed in the technical manuals offered by the equipment manufacturers. However, it has been a common discouraging experience that the user must request a copy of the appropriate manual at least a dozen times before one is provided. Even then, a manual will not supplant the careful pretesting of the answer sheet recommended earlier in this handbook.

A few suggestions for the format of the answer sheet are: (1) that it include provision for sample questions; (2) that the minimum number of columns be used for student number in order to reduce the number of candidate errors in recording this information; and (3) that when a computer is available, the office preprint the candidate's name, number, school, sex, and age on the answer sheet to further reduce the risk of errors. An example of a precoded answer sheet used in West Africa is shown in Figure 6.2, and an answer card used in Southern Africa in Figure 6.3.

The printing of the answer sheet should be done in the same color as the test booklet when possible. If the latter is red, the former should be red, and the directions should specify that the two items match in color. This procedure worked very well in West Africa for a number of years, until someone in the computer section decided that an all-blue answer sheet would look nicer and had half a million printed. Naturally, he forgot to change the directions in the test. As one might imagine, the error was spotted by the candidates who, puzzled by the instructions, were wildly and unsuccessfully trying to match answer sheet to test booklet.

One of the dangers in the reproduction process is that with all the rearranging and modifications on the test and answer sheet, the scoring key might become confused. The surest way to prevent an error is to produce a correct answer key at the time the test and the answer sheet are being printed.

Another comparatively easy way to avoid problems for the STO and to help the candidates at the same time is to provide any tables needed for computation or answering a question so candidates will not have to use their books of tables. Sometimes the pertinent sections of the tables can be reproduced along with

Figure 6.2. Precoded Answer Sheet

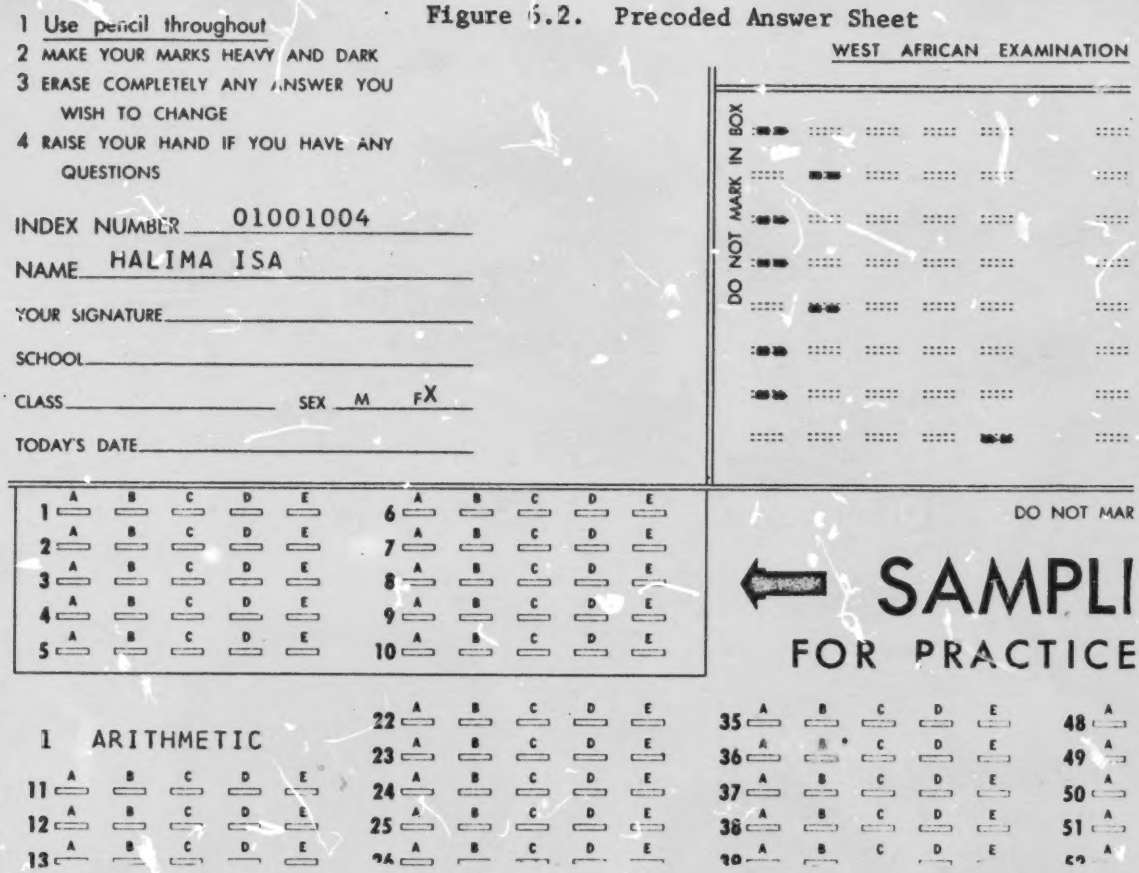
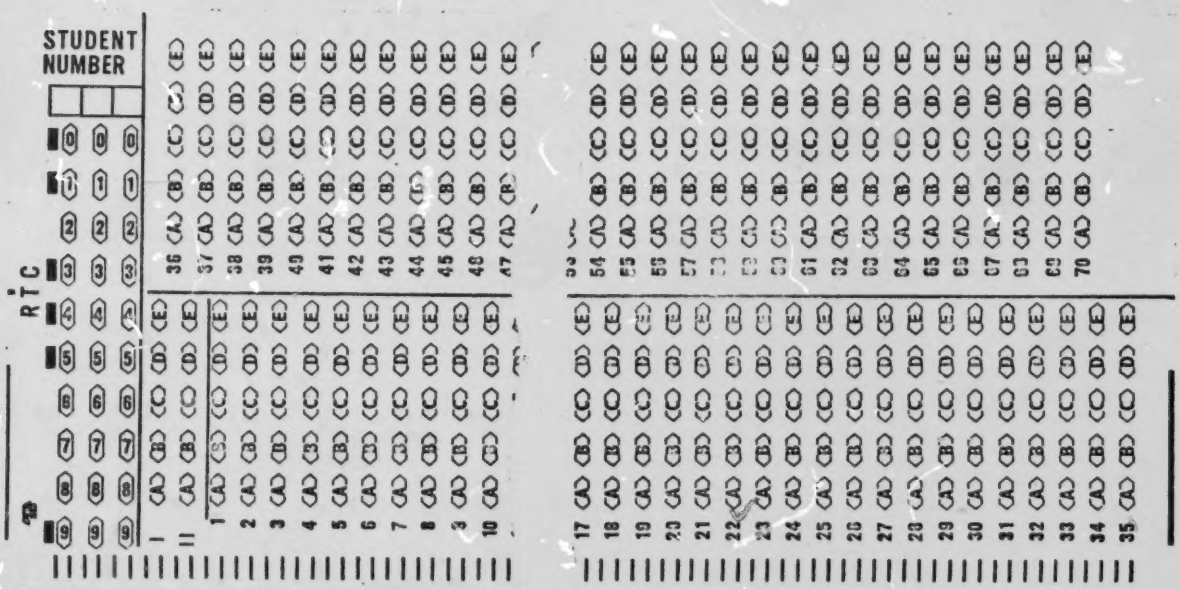


Figure 6.3 Answer card



the test booklets, perhaps on a page which would otherwise have been blank, and consequently ensure uniformity of source materials. Such a system also eliminates the possibility of candidates' cheating by writing answers in the margins of their books of tables.

The practice of providing the candidates with scratch paper for rough work differs widely from country to country. In some places, a candidate is told to do his work in the margins or on a blank page of the test booklet, which saves much administrative time and the cost of extra paper. In others, the candidate is given a separate sheet of paper. Obviously, if the STO desires to collect all the booklets, unmarked, for later sale, separate papers must be provided. Sometimes in the past, these booklets have been marked: "This paper is confidential and the property of the State Testing Office. It must be returned." Sometimes the papers have been collected and actually returned to the office, but more often thay have simply been ignored, because no large-scale examination is secure once it has taken place, and collection of the paper loses meaning in the light of this fact.

The security precautions are vital, however, in the printing of the papers. (General principles for printing security are given in the overall security section, page 9-16.) Perhaps some idea of the value set on these papers before the examination can be gauged by the fact that the regular price for illicit copies of the test in 1965 in West Africa was $140, with no questions asked of the seller. Because of this, printing the papers was very similar to printing money; in fact, in one country the office requested the government printers (the mint) to reproduce their papers. They refused, preferring the relative security of printing currency.

With any method of reproduction, proofreading is vital, and printing with the offset press is no exception. In preparation for using this method, the test is typed and the diagrams and drawings taken from textbooks are drawn on tracing paper and attached with rubber cement to the textual materials. This is the camera-ready copy, and is proofread thoroughly at least twice before it is considered ready for reproduction. Only after these careful proofreadings does the copy leave the confines of the office.

In West Africa, a senior officer was commissioned to take the copy to the printer and to wait while the negatives were developed. When they were dry, he watched while they were stripped and all was made ready for the plate making. At that point, he took the copy, the negatives, and any other

test-related material back to the STO, since printing could not be begun on that day.

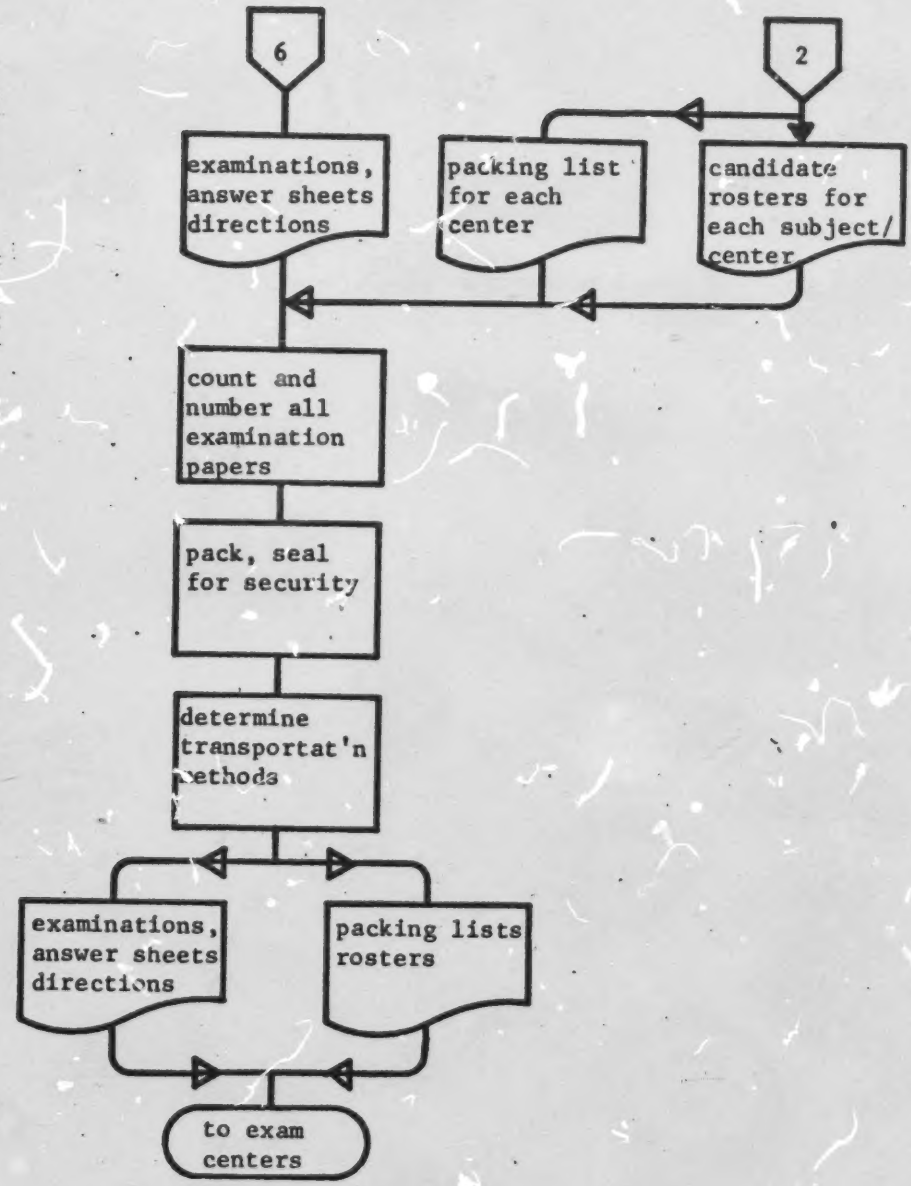
The following morning at least two senior officers returned to the printer where everything was ready for a long day of printing. The plates were made from the negatives which the senior officers brought back with them, and were placed directly on the press. All doors were locked, and the press, a web offset with binding attachment, proceeded to turn out 20,000 copies of the booklet per hour. Every single piece of waste paper, no matter how small, was placed in a large sack (with triple stitching at the seams) to be taken back to the STO by the staff members, shredded, and burned. At the end of a long working day, the printing of 200,000 test booklets was completed, and all traces of the job removed.

Another, and considerable, advantage of the offset method should be apparent from this illustration. With the letterpress, there is no guarantee of security while the copy is set in type, while galley proofs are being made, or while the papers are being printed, since it is impractical for a senior officer to stay at the printers for the month that this process typically takes. It is, however, imminently practical for him to stay through the offset process, which totals about a day and a half.

During five years of following this scenario, there were no detected instances of a security leak. The printing was done only two weeks before the distribution of the papers was scheduled to begin, so no large supply of confidential material had to be stored and guarded for a long period of time. The printing was done on inexpensive newsprint, in two colors, at less than four cents per copy.

The appearance of the final copy was reasonably satisfying aesthetically, and the cost was minimal. After the booklets were used, they were destroyed. Some sort of compromise of a similar type must be made in every case of reproduction of tests for a large-scale examination.

7. Distribution/Administration



DISTRIBUTION/ADMINISTRATION

Even if the final form of the test is a model of precise, well-constructed, worthwhile, technically perfect questions, but the administration is faulty, the results are useless. No amount of extra care in the formulation of objectives, in the editing of items, in the adequacy of the pretesting, can compensate for any misunderstanding of directions or errors in indicating answers.

As stressed earlier, if the various aspects of distribution and administration are pretested along with the questions themselves, the number of errors will be drastically reduced. Other considerations are discussed below, and many of the security aspects are taken up in a subsequent chapter on overall security.

The number of candidates expected at each testing center should be known as a result of the applicant registration. The materials--test booklets, answer sheets, directions, scratch paper, return envelopes--can then be counted out for each center. (See Figure 2.3 for an example of a packing list.) Depending on the time and personnel available, the materials can be counted at the time of shipping or they can have been precounted into some arbitrary-sized bundles well in advance. The latter method allows the counting to be spread out over a longer interval and so balances the work load somewhat. If testing materials are placed in bundles of 5, 10, 25, 50, and 100, and if each bundle is recounted, and then wrapped, it remains only to choose the appropriate number of bundles and to place them in a shipping carton at test time. By using the computer to print out the packing lists, the process can be simplified and speeded up considerably. Each list specifies the name of the testing center, number of each size package included, number of instruction manuals, and even the appropriate mode of transport.

Precounting allows the bulk of the simple repetitious work to be done well in advance of the last-minute rush of shipping. If examinations have been printed in another country, this method provides enough lead time to order additional tests if they are needed. One serious breach of security occurred when, during a late packing, an insufficiency of booklets was discovered. Additional copies were airfreighted from London and the packing proceeded hastily and inaccurately. Some packages contained enough surplus papers that many supervisors were tempted to remove them, and some succumbed to the temptation. To the cost of the extra airfreighting charges were added

the costs of poor publicity and an additional testing, well in excess of $20,000.

When counting takes place at time of shipping, security is threatened in another way. The candidates' anxiety is transmitted to the office employees at a time when the office employees have access to the actual tests. If the precounting takes place well in advance, before the test anxiety has built up, the security is tighter. In addition, test papers can be counted twice, and even weighed to make certain that no count is erroneous. Then the bundle of papers can be sealed in a plastic sheath and stored until the main packing dates.

Some offices, hypersensitive to the need for security, also stamp every examination paper with a number. They contend that the added effort of using an automatic stamper which increments after every impression pays off in accuracy and ease of accountability. All through the process, including the shipping, papers can be traced very easily. If any papers are found on the black market, the number will give a clue to the source.

Details of the packing procedures and some of the transportation considerations are given in the chapter on overall security.

The most secure modes of transportation differ from country to country. In some places, the rail system is comprehensive and reasonably safe (despite which one office always hires an armed professional guard to ride with every shipment), and in others the postal system is preferable. In an emergency in Western Samoa once, a staff member gave a package to a stranger just boarding a bus and asked him to see that it reached its destination. Someone else delivered the package intact in record time; no one knows how many individuals handled it along the way. No matter what method of transportation is chosen, the staff always say prayers, throw bones, sprinkle ashes to the wind, beat drums, offer up chickens--and most of the time it works.

Once the supervisor receives the papers, another combination of worries takes prominence: will the examination be given in a standardized way, with the proper time limits, and so that every candidate has a chance to do his best? The STO worries less if, whenever possible, they use experienced examiners. Sometimes this means using someone as an invigilator for a year or two before making him a supervisor. Since an invigilator's responsibility is to see that every candidate is filling in his number and answers correctly, and that no one is either giving help or receiving aid, and since he is performing under supervision, he can be considered to be

properly trained. If it becomes necessary to use inexperienced personnel, the use of students from universities and teacher training colleges is preferred to experienced teachers. When students have been used, they have always been ready to learn, and willing to follow the directions exactly. In contrast, the teachers have thought they knew everything already, and often could not be bothered with someone else's directions; after all, they had been giving tests for years and years.

As will be reiterated later, there is merit in frequent, unannounced site inspections during the actual testing. Every single senior staff member should visit some center at every major examination and record his impressions. The insights into the problems of administration gained in this way are invaluable. Other kinds of insights occur as well. This year in one country a staff member found candidates entering the examination hall with their answer sheets already completed.

An item which the visiting staff member should always check is the certificate of examination conduct (see Figure 7.1). On it a supervisor records details of the test administration, irregularities, etc. An easy method of keeping the packing materials and supervisor record together is to print the certificate on the side of the envelope used for answer sheets. When the certificate is returned, an inspection of the envelope and certificate can give clues to how well the materials survived the shipping and how closely the actual administration paralleled the ideal. If the actual times begun and ended deviate markedly from the scheduled times, changes can be made in the directions or procedures to bring reality a bit closer next time.

An example of a change brought about in this way might be of interest. When the examination was scheduled to begin at 07.30 and the certificates from one section of the country reported 09.30, an inventigation was made. The examination took place on a Monday morning, and all the candidates had had to return to school from villages at considerable distance away. Since they were unwilling to travel until daylight, they could not arrive until almost 09.00. The starting time was rescheduled to 09.30; the day could have been changed to Tuesday as well.

Often the supervisor and invigilator must travel to reach the center too. If so, allowance must be made for the vagaries of public transport and the conditions of the roads. A decision must also be taken about reimbursement for transportation expenses.

Figure 7.1 Certification of Examination Conduct

1. I conducted the ABC Examination on the following time schedule:

Paper:	Time limits:	Actual time begun:	Actual time ended:	Actual testing time:
Test A	1 hr. 5 min.			
Test B	1 hr. 10 min.			
Test C	1 hr. 45 min.			

2. I wrote "absent" next to the name of every candidate on the attendance register who did not appear for the examination.

3. I seated the candidates according to the seating plan which is attached, and placed them so as to prevent them copying from each other.

4. I found no one cheating with the following exceptions:

_____

(write 'none' if no cheating was found; otherwise attach an explanation)

5. I read the instructions in the manual, and followed them throughout the examination.

I hereby certify that these tests were administered in exact accordance with instructions and that the facts stated above are true. I also certify that I have withheld no information regarding irregularities which reflect on the conduct of the examination.

Dated this _____day of _____,
in the year _____.

signature:_____
(supervisor)

signature:_____
(invigilator)

supervisor:_____
(in block capitals)

For certain types of elaborate examination instructions, (generally avoided in large-scale examinations), the supervisor will have to be provided with visual aids. One method adopted in the development of tests in West Africa was the production of display answer sheets, question papers, and other explanatory materials on heavy plastic. The supervisor, using a grease pencil, could demonstrate to the candidates the correct methods of answering and marking the answers on the answer sheet. The grease pencil could be wiped off with a rag and the materials re-used. If a less expensive, less durable display were desired, heavy bristol board could be printed and then covered with a layer of PVC plastic, the type that already has an adhesive backing. For the more complicated types of directions, such an action is necessary since it will not suffice for the supervisor merely to use a blackboard (and some examination halls do not even have blackboards).

When the examination has been completed, the supervisor must finish filling out the supervisor certificate, package the materials as outlined in his instructions, and return the papers as directed. If transportation is not a problem, all the materials can be returned at once; otherwise the answer sheets can be sent back immediately by the fastest possible means and the remaining materials sent by the cheapest possible means. Whatever the method, the STO should provide prelabelled, complete repackaging supplies for each testing center unless the office is prepared to wait for a few months while a shipment is traced or the candidates retested. If the test has been scheduled for any day except Friday, the supervisor can ship the materials the following day rather than keeping them over the weekend.

When all the test booklets are returned, they can be inspected and those which are unmarked by the candidates can be sold to future candidates who want an old examination on which to practice. No matter how careful a STO may be, seldom if ever is a large-scale examination secure after it is once given, especially if it is administered by other than STO personnel. Thus a liability can be turned into an asset if the used papers are sold. If the sale is widely advertised, candidates will not be as willing to risk getting caught removing examinations from the testing hall. But more important for many offices is the added income this provides: a booklet which cost about four cents can be sold for fifteen. If the STO prefers not to become so commercial, permission to distribute and sell the papers can be given to a local publisher or bookseller with a portion of the profits to be returned to the testing office.

In those cases where the tests are *not* to be sold, the simplest, shortest, least expensive method is merely to let the candidates keep the booklets!

Specific instructions must be given to the supervisors concerning timing. Lack of adherence to specified time limits is probably the most common error in the area of administration, and one which arises as a result of a number of factors: a supervisor's lack of a reliable watch or other timepiece, difficulty in calculating the amount of elapsed time, tendency to give candidates extra time if they are not doing well, and failure to read the timing directions carefully. Is there an office which makes a practice of ascertaining if supervisors even have a watch, let alone if it is running and accurate?

The directions may specify that the supervisors should reset their watches to some arbitrary starting time (e.g., 08.00) as the candidates begin the actual test. This practice would counteract the difficulties of the supervisors in calculating stopping times. If the supervisors do not reset their watches but mark down the actual starting time, hopefully they can calculate the proper stopping time.

The problem of the generosity of some supervisors in allowing extra time for their candidates is one which cannot be solved easily; perhaps the only way is to make spot checks while the examination is in progress. In those countries where the appointment of supervisors is made on a political basis, the problem is particularly acute since their continued appointment depends on the good will of their constituency. (One of these appointees in a recent examination was illiterate as was unfortunately found out too late.)

If complete schedules are given, care must be taken to be sure that supervisors know that they can depart from the times given with no fear of penalty. For instance, if the candidates are to be seated at 07.33, booklets distributed at 07.38, test begun at 08.00 exactly, etc., and the poor supervisor finds that he actually began the test at 08.12, he should be able to modify the remaining schedule accordingly. Otherwise his tendency is to falsify the record or to hasten through later directions in the hope of making up the lost time.

The easiest way to avoid this entire problem is to give abundant time for every test; i.e., transform it from a speed test to a power test.

The suggested schedule should include provision for breaks, and for the length of time candidates can go without such niceties as food, water, and the use of lavatories. Such an obvious point is mentioned only because it has sometimes been overlooked. In regard to time, the limits on the sample exercises should be proportional to those used during the actual test. A sample set of directions for administration begins on page 7-9.

Two additional security precautions can be mentioned here: edge seals and plastic envelopes. The edge seals are specially printed, gummed circles of paper which are used to seal the edges of the test booklets. When the candidate is ready to begin, he inserts the eraser end of his pencil under the seal and breaks it; any broken seal is proof that the booklet has been tampered with. These can be printed to order almost anywhere.

In the earlier paragraphs on shipping, the use of plastic envelopes for the test booklets was mentioned. Plastic sheathing, available for rapid wrapping of small parcels, comes in rolls of various widths. The booklets are inserted, a combination cutter/sealer knife is applied, and the package is sealed against moisture and people.

Finally, specific directions should be given to the Special Education Officers or to the supervisors about what to do in various emergencies such as the following:

1. they have insufficient papers.
2. they spot typographical errors.
3. their packages have been tampered with.
4. their candidates do not have pencils.
5. a candidate becomes too sick to continue.
6. they collect more test papers than they distributed.

If no emergencies occur, if all directions have been carefully observed, and if all papers are promptly returned, one can go confidently into the next phase.

Instructions for Administering the ABC Series

A. Introduction

Supervisors and invigilators should read this manual at least one week before the actual testing takes place.

The major responsibility of the supervisors and invigilators of an examination is to help the candidate do his best work. They do this by carefully controlling the conditions under which the candidate performs, and by making certain that he understands exactly what he is to do. As far as the answering of the actual examination questions is concerned, however, they must give no help whatsoever. Once the candidate has begun answering these questions, the supervisor and invigilator must repeat no instructions, make no explanations of difficult words, phrases, or concepts contained in the questions, indicate in no way whether the candidate has answered a question correctly or incorrectly. In short, they must never assist or encourage the candidate toward the choice of an answer. They are free to explain misprints *only*.

B. Examination

In the examination, there are three papers: A, B, and C. Each paper consists of a number of questions, each followed by several possible answers, only one of which is correct. The candidate must indicate the one answer he considers to be correct by completely filling in the space on the answer sheet which has the letter of the answer he thinks is right. If, for example, in question 1 the correct answer is B, this would be recorded as follows:

1. The first month of the year is

A. December
B. January
C. September
D. July
E. November

The candidate must blacken the box completely. He must not go outside the box. For each of the three objective test papers, he records his answers on a separate answer sheet. Supervisors and invigilators should make sure that all candidates know how to record their answers on the separate answer sheets before the examination begins. This should not be difficult because specimen questions, and answer sheets with detailed instructions on their use, have already been sent to candidates.

C. Supervisor Instructions

Each supervisor will receive an attendance list, certification form,

the A, B, and C test papers, answer sheets, and other examination materials including brown paper, string, sealing wax, and large pre-addressed and stamped envelopes from the Special Education Officer, probably the day before the examination. The sealed envelopes containing the question papers *must not* be opened before the actual time of the examination, and only under the conditions outlined below.

If the materials are given to the supervisor on the day before the examination, he must keep them in complete security. Where no safe is available, they should be kept in a locked box to which he alone has access. Supervisors are reminded that as soon as they have received and signed a receipt for the examination materials, they are responsible for their security. This means that if anyone sees the examination for any reason by any method before it is officially distributed to the candidates, the supervisor is at fault. Such instances will be investigated by the Criminal Investigation Division and criminal charges preferred.
[Students from those schools where cheating has occurred will not be allowed to take the examination in following years.]

The supervisor should be sure that the following materials are available at the testing hall on the day of testing:

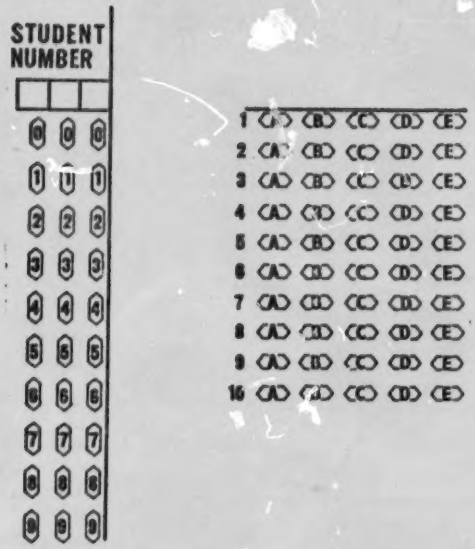
1. the Supervisor's Instructions (this booklet).
2. the examination materials.
3. a supply of extra pencils (with erasers).
4. a timer or reliable time piece.
5. a blackboard and chalk.
6. a pencil sharpener or razor blades.

D. Before the Examination

The supervisor and invigilators should arrive at the testing hall at least 30 minutes before testing is scheduled to being to arrange it properly, locate lavatories and other facilities, and to draw the following diagram of a section of the answer sheet on the blackboard to be used to demonstrate to the candidates how to mark the answer sheets:

(diagram is at top of next page)

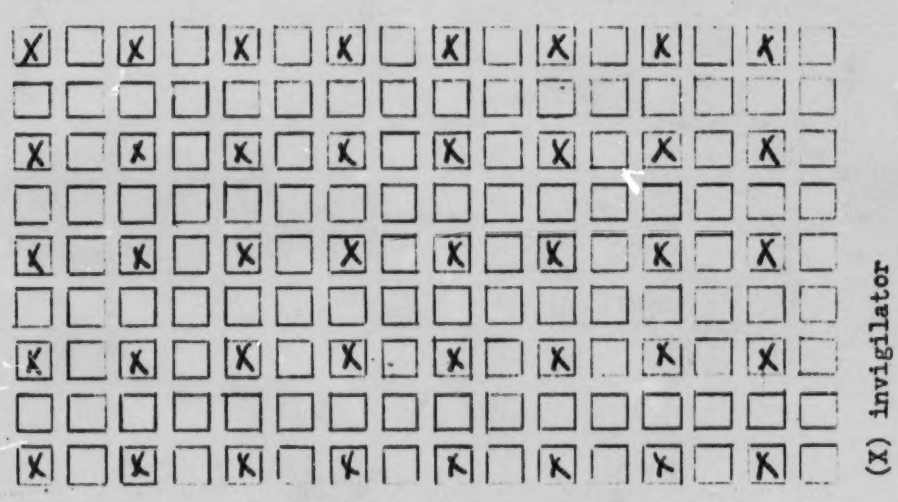
A separate desk should be provided for each person who is expected to take the examination. Such physical separation reduces the temptation to copy or to be distracted by motion at the side. Ample space should be provided between rows of seats to permit distribution of test materials to



each candidate individually, and to permit invigilators to observe candidates' work and to give individual help when it is needed.

The supervisor's station should be at the front of the group, and a table or desk should be available for laying out testing materials. All candidates' desks must face in the direction of the supervisor's station. A typical hall layout for 40 candidates appears in Figure 7.2.

Figure 7.2 Typical examination hall layout for 40 candidates.



E. Admission of Candidates in the Testing Hall

When the scheduled time arrives, supervisors open the hall door and admit candidates. As each one enters, his admission notice is checked and his name marked off on the attendance register. As they enter, candidates

should be directed to seats on alternate sides of the room in order to separate friends who come in together. *UNDER NO CIRCUMSTANCES SHOULD CANDIDATES BY ALLOWED TO CHOOSE THEIR OWN SEATS.* The rows closest to the supervisor should be filled first, and every designated seat should be occupied within the rows used.

Some candidates may have incorrect admission notices or be carrying letters or telegrams of authorization. After all candidates who present valid admission notices and who are listed on the attendance register are admitted, other examinees should be admitted in the following order, as long as surplus testing supplies last (the quantities are listed on the packing list):

1. Those who have letters or wires of authorization for that center, although they may not be listed on the attendance register.

2. Those who do *not* have valid admission notices or authorizations, but who are listed on the attendance register.

3. Those who have an admission notice for another center.

The names of any candidates admitted under 1 or 3 above must be added to the attendance register. It may be necessary to remind these candidates of their examination numbers at a later stage.

[An alternate procedure is as follows: Using the attendance register, the seating of the candidates should be planned and their index numbers written on their desks with chalk or stuck on with a label. Then as the candidates enter, they are directed to the desks bearing their numbers. Some supervisors object to the gaps left by absentees when this system is used, but like the ease of placement.]

The doors of the room are to be closed at the hour previously announced. Any candidate who arrives late may be admitted at the discretion of the supervisor, so long as the actual test administration has not begun and the admittance of this candidate will not disturb the others. Under ordinary circumstances, no one should be admitted after the actual testing has begun. However, if the supervisor decides to admit someone who has been delayed by an unusual circumstance, he must assume the responsibility of giving him the full testing time.

When all examinees are seated, the supervisor should make an exact count of those actually present, and compare this count with the number of tallies on the attendance register. The register should be corrected if necessary.

F. Opening the Sealed Papers

The envelopes containing the examination papers for the subject stated on the examination timetable should then be opened by the invigilator in the presence of the candidates and the supervisor. The supervisor and invigilator will inspect the materials and compare the quantitites received with the quantitites on the packing list. Any differences between the materials received and those listed must be immediately noted. Both the supervisor and the invigilator should then complete and sign the certificate.

FROM THIS MOMENT UNTIL THE END OF THE TESTING SESSION, NO PERSON, EXCEPT THE CANDIDATES, THE SUPERVISOR, THE INVIGILATOR(S), AND MINISTRY OF EDUCATION OFFICIALS, IS PERMITTED TO ENTER THE EXAMINATION ROOM UNDER ANY CIRCUMSTANCES.

G. Conducting the Examination

It is very important that all supervisors give their tests in the same way. That is, they should read the instructions to the candidates exactly as they are given in this manual. While it may sometimes be necessary to answer questions or even to add some instructions when candidates clearly do not understand, the instructions in the manual should be given exactly as written. It is advisable for the supervisor to read through these instructions several times before the testing session so that he is thoroughly familiar with all instructions which he must read word for word. Throughout this manual, the paragraphs which *must* be read exactly are enclosed in boxes. Whatever is inside the boxes must be read word for word. Any words or phrases in the boxes which are enclosed in parentheses, however, are not to be read; these are additional instructions for administrators.

As soon as he has completed the above procedures, the supervisor is ready to begin conducting the examination.

He says:

Listen carefully to everything I say. The tests you will take today may be different from any tests you have ever taken before. Before you begin each paper, I shall explain to you how to do the paper and how to mark the answers. If you listen to me, you will know exactly what to do. You will also have time to practice the tests before we begin.

The only things you will need for these tests are your pencils and erasers. Put everything else under your seats. (PAUSE WHILE CANDIDATES CLEAR THEIR DESKS.) Everyone, hold up your pencils.

The supervisor should be sure that everyone has a pencil. All candidates were instructed to bring two pencils with them to the examination, but the supervisor may have to provide some to those who have forgotten. He should also give every pupil two sheets of paper for rough work [or tell the candidates to do their rough work in the margin of the test booklet.]

The invigilator should then hand each candidate his test booklet individually. He must be absolutely certain that each person receives only one test booklet. After distribution [or after the beginning of testing], the supervisor must check that the number of papers given out, plus the number remaining, equals the number he was supplied by the STO. He must be able to account for every paper.

He then says:

Turn over your test booklet. Everyone should have a booklet which looks like this one in my hand. (HOLD UP A TEST BOOKLET) You should also have a separate paper which looks like this one in my hand. (HOLD UP A SEPARATE ANSWER SHEET). Who does not have a test booklet and an answer sheet like these? (PAUSE)

The tests you are to take are in three parts and the total time you will work on these tests will be about three hours. The tests may be different from those you have taken before. They are a new kind which is called "objective." In this new type of test, you are given five possible answers to each question. You must decide which one of the five answer is the correct one. If none of the answers appears to be the correct one, you should choose the BEST one. Only one answer will be marked correct, so you must choose ONLY ONE for each question. If you cannot decide which is the correct, or best, answer of those given, you should leave the problem and go on to the next question.

You are NOT to mark your answers to the questions in the test booklet. You will mark all of your answers on this separate paper. (AGAIN HOLD UP THE ANSWER SHEET.) In a few minutes I will show you how to make your marks on this sheet.

Only pencils must be used for writing or marking on this answer paper. This is because your marks will be read by a machine. This machine can read only pencil marks. But it will read any pencil mark you put on the paper, and dark fingerprints and smudges as well. So if you make a mistake and want to change a mark, you must erase carefully and completely. (PAUSE.)

Now look at your answer paper. Notice that it is printed on both sides. Turn your paper back to the front side. (PAUSE)
[Turn the paper like this (DEMONSTRATE) and find where it says Surname. I want you to write your surname or your father's name on this line.(POINT) Everybody write your surname. (PAUSE). Just under your surname there is a place to write your other names. Write your other names.(PAUSE).Now write the name of your school. (PAUSE). Write your class and your age. (PAUSE) Have you finished?(PAUSE). If you are a girl, put a circle around the letter "F"; if you are a boy, put a circle around the letter "M." (PAUSE) Put down today's date, which is _____ and sign your name. (WRITE THE DATE ON THE BLACKBOARD.) When you have finished, put your pencils down.

(Throughout these instructions invigilators should walk slowly through the testing room, watching as the candidates try to follow the directions. They should help any candidate who is having difficulty.)]

You have all been given an examination number and it is written on your admission notice. If this number is not entered correctly on your answer sheet, your examination will not be marked. If you cannot find your number, please raise your hand. (PAUSE). I want you to mark your examination number here.(POINT TO POSITION ON ANSWER PAPER.) Everybody look up here. (POINT TO THE DIAGRAM ON THE BLACKBOARD). I am going to show you how to mark my examination number on my answer paper. My examination number is 852 (WRITE 852 ON THE BLACKBOARD.) So I write this number in the first three boxes, like this. See, I write the first number--8--in the first box. (WRITE IN THE NUMBER ON THE BLACKBOARD DIAGRAM). I write the second number--5--in the second box, like this. (DEMONSTRATE). I write the third number--2--in the third box, like this. (DEMONSTRATE). See, I have written my examination number--8,5,2 in the first three boxes. Everybody write YOUR OWN examination number in the first three boxes on your answer sheet. (PAUSE).If you do not understand how to write your examination number in the boxes, please raise your hand. (PAUSE).

Invigilators should walk around the room, checking to make sure that every candidate writes his number correctly in the proper spaces.

Now I am going to mark the numbers in the spaces below the boxes. First I go down the first column and find the small box with the number 8. Then I take my pencil and blacken out the small box like this. (DEMONSTRATE.) I make my mark heavy and black, and completely fill the box, but I do not

go outside the box. (PAUSE). Everybody find the small box in the first column which is the same as the number YOU wrote in the first larger box. When you find the right box, blacken it out with your pencil. (PAUSE.) Now look at the number I wrote in the second box. I now go down the SECOND column and find the small box numbered 5 and I blacken it out with my pencil like this. (DEMONSTRATE). Everybody find your second number in your second column and blacken out the small square. (PAUSE). Now find the small box in the third column numbered the same as your third square and blacken it. (PAUSE AND BLACKEN PROPER BOX ON BLACKBOARD.) Has everyone blackened his third number? (PAUSE).

Everybody turn your answer paper over and find the place for marking your examination number. (DEMONSTRATE, PAUSE). If you cannot find the examination number spaces, will you please raise your hand. (PAUSE) Write your examination number in the three boxes just like you did on the front side of the answer paper. Everybody write your number. (PAUSE). Now find the correctly numbered box for each column and blacken it out. (PAUSE). Have all of you marked in your examination number? If you are having trouble, raise your hand. (PAUSE) Now everyone write your examination number in the upper right corner of your test booklet. Also write your name just under your number in the space provided. (PAUSE UNTIL EVERYONE IS FINISHED) Everybody, pencils down.

While the supervisor is reading the above instructions, invigilators should check each candidate's marking, and provide individual instruction or assistance as needed. The most common error is for a candidate to blacken more than one number in a single column.

There must be no talking until the examination is over. You should read the instructions at the beginning of the paper very carefully while you are waiting to be told to begin. Do not read the questions themselves until I tell you to do so. Now look at the paper entitled INSTRUCTIONS. Read this to yourself while I read it aloud.

1. Number one. You must only use pencil for this examination. The answer sheet is to be marked by a machine which can only read pencil marks. Do NOT use a ball point pen; do NOT use an ink pen; do NOT make any mark which cannot be rubbed out.

2. Number two. When you have chosen the answer to a question, blacken the box which has the same letter as the letter of the answer you have

chosen.

3. Number three. Examples have been given to help you. Read them carefully before doing the questions and listen carefully to the instructions

4. Number four. If you wish to change any answer you have chosen, use a rubber or an eraser to rub out the mark completely, and then make a new mark. If you do not erase completely, the machine may mark you wrong.

5. Number five. When you are told to begin, work as fast and as accurately as you can. If you cannot answer any question, do not spend too much time on it; go on to the next question.

6. Number isix. Make no marks on the question paper at all. Do your rough work on the sheets provided, NOT on the question paper.

7. Number seven. Do not ask for further explanations after the examination has begun. During the examination, the supervisor is allowed to answer questions about misprints ONLY.

8. Number eight. This question paper will be collected by the supervisor. It is the property of the State Testing Office and must NOT be removed from the examination hall. No candidate will be allowed to leave the examination hall until the supervisor has collected ALL copies of the paper.

9. Number nine. There is absolutely no smoking allowed in the examination hall.

Are there any questions? (PAUSE)

Now, insert the eraser end of your pencil under the seal on your test booklet and break the seal.

Turn to the next page for the "A" test practice problems. I will discuss questions one to five with you and show you how to mark your answers on the answer sheet. You must then answer questions six to ten on your own and I will check the answers with you. Everybody read questions one to ten.

(AFTER TWO MINUTES, THE SUPERVISOR STARTS TO READ SLOWLY BUT LOUDLY THE FOLLOWING:)

Do not turn over the page until I tell you to do so.

Question 1: 5 + 3 + 4. The answer is 12, which is letter C. Therefore, on the answer sheet we blacken box C for question 1, as shown in your booklet Now, on your answer sheet, blacken box C for number 1. (INVIGILATORS GO AROUND TO SEE THAT THIS IS CORRECTLY DONE FOR EACH EXAMPLE. MOST

INEXPERIENCED CANDIDATES WILL NOT VOLUNTARILY ADMIT THAT THEY ARE CONFUSED.)

Question 2: 7 + 4 - 3. The answer is 8, letter B. Now, on your answer sheet blacken box B for number 2.

Question 3: 207 + 87. The answer is 294, letter D. Now, on your answer sheet blacken box D for number 3.

Question 4: 0.5 X 0.2. The answer is .10, letter A. Now, on your answer sheet blacken box A for number 4.

Question 5: Find the average or 8, 4, and 6.

$\frac{8 + 4 + 6}{3} = \frac{18}{3} = 6$ The answer is 6, letter B. Now, on your answer sheet blacken box B for number 5.

Now answer questions 6 to 10 on your own and I will check the answers with you. You may do your work in your head or on the rough paper. You have five minutes. Begin. (INVIGILATORS GO AROUND TO SEE THAT CANDIDATES MARK THEIR ANSWERS PROPERLY).

(AFTER EXACTLY FIVE MINUTES.) STOP, make your corrections as I give you the answers. Remember to erase an answer completely when you wish to change it.

Question 6: The answer to question 6 is X, so we blacken box D.

Question 7: The answer to question 7 is 3.1416 so we blacken box B.

Question 8: The answer to question 8 is 22, so we blacken box A.

Question 9: The answer to question 9 is 0.45, so we blacken box C

Question 10: The answer to question 10 is "None of these" so we blacken box E. (PAUSE)

Are there any questions? (AFTER MAKING ANY NECESSARY EXPLANATIONS, SAY:)

You have 65 minutes to answer questions 1 to 60. Turn over the page. BEGIN.

Set your watch to exactly 08.00. If anything is found to be wrong with the test materials, give the candidate new materials and attach an explanation to the examination report.

After exactly 65 minutes, when your watch should say 09.05, say:

STOP. Pencils up. Close your test booklets and put your pencils down. I am now going to collect the test booklets and then the answer sheets.

Invigilators collect *first* the test booklets and count them. When this number agrees with the number given out, the answer sheets are collected and also counted. When the number of answer sheets collected agrees with

the number given out, candidates are sent out for 30 minutes' break (or sent home if they are finished for the day.) If a break is given, be certain to guard the examination materials closely during it.

:
: (remaining instructions for each test paper)
:

H. Concluding the Examination

After all the examination booklets and answer sheets have been collected and counted, the supervisor should check that the numbers of each agree with the packing list. He should then repackage all material in the pre-addressed container provided. The invigilator should then seal the envelope in the presence of the supervisor, after which they should complete the certificate together and sign their names across the joints of the envelope as a check against scripts being tampered with. Finally, they should use the sealing wax across the signatures on the envelope, and then insert the envelope in the larger stamped and addressed package for return to the State Testing Office.

All question papers must be returned to the STO, including those issued to candidates. Supervisors and invigilators will be held responsible for any shortage of question papers, and marking of papers may be withheld until all question papers and answer sheets issued are received from centers.

I. General Suggestions

Giving Instructions: All instructions should be delivered by the supervisor only. He should speak loudly, clearly, and slowly enough that every person taking the tests can hear and understand every word he says. His manner, as well as his voice, should be friendly and encouraging. The importance of the supervisor's attitude cannot be overestimated; it can either help a student do his best, or cause him to do less well than he can. The supervisor's aim, therefore, must be to be regarded by the candidates as a friendly person who sincerely wants to help them succeed on the tests.

Timing: Each of the papers has a time limit which must be followed exactly. The supervisor and invigilators should watch the candidates closely throughout the testing period to be sure that they are working on the correct paper and that they have not gone ahead or turned back to other parts of the examination.

No matter how clear the directions are for a test, one or more candidates are sure to misunderstand or misread them. They may fail to mark their answers

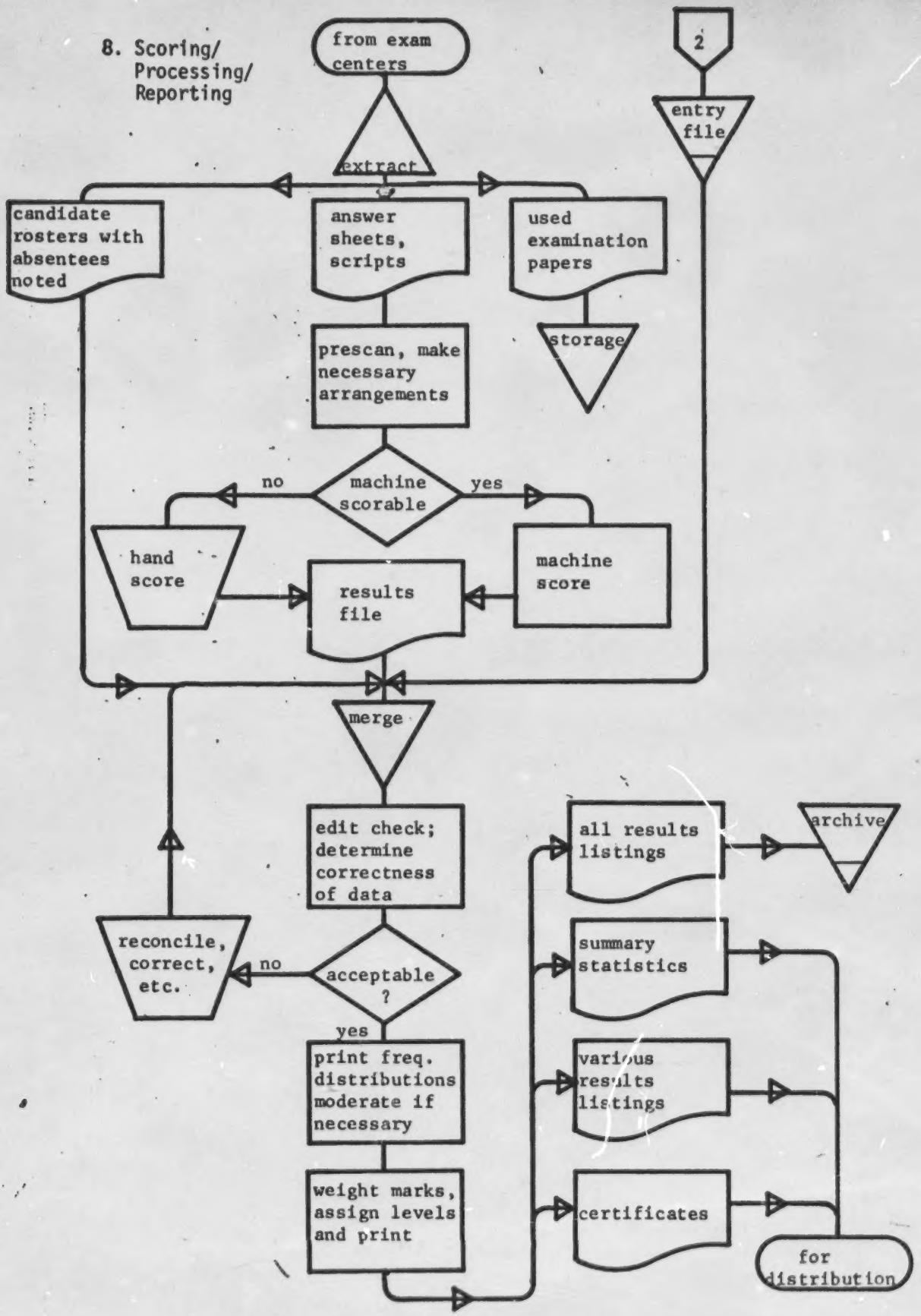
on the answer sheet as directed; they may fail to continue a test in the belief that they have finished; or they may find that they have marked their answers on the wrong part of the answer sheet. *Time lost by a candidate through his own failure to understand or follow directions cannot be made up or taken into account in marking.*

J. Invigilating

When tests are administered to a large group, one invigilator is needed for every 25 examinees beyond the first 25. No invigilator is needed if there are 25 or fewer candidates. When there is more than one invigilator, the supervisor should assign a definite section of seats to each one before the testing. Invigilators should help distribute and collect test materials, observe candidates' work during the practice periods to see that they are marking their papers properly and following instructions, and patrol their sections during the testing. The invigilator will have to keep moving during the testing to do his job effectively. In addition, if he pauses too long behind a candidate, he may disturb or embarrass him. If a candidate is permitted to leave the room at any time, his invigilator must make sure that his test books are left closed on his desk with answer sheets inserted inside the front cover. If two or more candidates leave the room at the same time, the invigilator should go with them; under no circumstances, however, should the examination room be left unattended. No extra testing time may be allowed to make up for a routine absence, such as a trip to the lavatory. *No candidate is allowed to leave the examination hall during the last ten minutes of the time allotted except in an emergency.*

Cheating is one of the most serious problems that arises in any examination. The supervisor and invigilators together, however, can do much to prevent and discourage acts of dishonesty before they occur. Copying can be prevented by physical separation, and by watchfulness on the part of the invigilators. If an invigilator suspects that one candidate is copying a neighbor's paper, he should change his seat immediately. If, however, the invigilator is convinced beyond a reasonable doubt that a candidate is giving or receiving assistance during the examination, he must collect the candidate's test booklet and answer sheet, and send the candidate from the room. The invigilator should explain to him that because he broke the rules, he must leave the room and cannot return. The word "cheating" should then be written across the face of the answer sheet and an explanation of the circumstances attached.

8. Scoring/ Processing/ Reporting



SCORING/PROCESSING/REPORTING

The processing steps from the receipt of the answer sheets to the issuing of the final results are shown on the flow chart, page 8-1. The documents are marked, after which the scores are checked for credibility, combined with the master list of candidate scores, and statistically adjusted. The adjusted scores are used to produce listings, certificates, and summary statistics.

One step which is often ignored despite its value is a general pre-scanning of the answer documents. The elaborateness of this process may vary from office to office. It can be anything from a casual glance through a few answer sheets, to a careful scrutiny and premarking of every single sheet. The general raison d'être is to assist the office in anticipating and preparing for the peculiarities of this particular batch of documents.

For example, in a recent scoring exercise the STO prescanned about two per cent of the answer cards and discovered that candidates from the far reaches of the Kalahari were often erasing their answers insufficiently. When they wanted to change their answers, they rubbed the card with some unknown substance which only imperceptibly lightened the marks. Consequently the office took two steps immediately: they brought in some teachers to assist in hand marking the troublesome cards, and they programmed the computer to identify any card with more than one inadequately erased mark. Later the administrative directions were changed in an attempt to prevent a recurrence.

One procedure which works well for hand scoring is for the staff to prescan the answer sheets and draw a red line through any question which has more than one answer option marked. The answer sheets can be counted and checked against the number of candidates present. In case of a discrepancy, the long process of reconciling the difference may be carried out while the scoring proceeds.

The actual scoring can be done with a fan or accordion key, a punched scoring stencil, self-scoring answer sheets, or machine-scored sheets. The fan or accordion key is simply a piece of paper with the correct answers marked in columns. Then the key is folded back and forth so a single column can be lined up with the candidate's answers. This works best for smaller examinations of perhaps up to five hundred answer sheets.

For larger examinations, typical offices will probably find the

punched scoring stencil the simplest and least expensive. A sheet of heavy paper (e.g., bristol board) the same size as the answer sheet has holes cut in it so the candidate's marks can show through. If holes are cut only for the correct answers, the scorer can simply count the number of marks which show through. Of course, he would not count the answers with a red mark through them since these have been identified as incorrect in the prescanning phase. Green sheets are usually preferred as they cause less eyestrain.

Two modifications to the usual punched scoring stencil procedure are recommended: The first is to use sheets of plastic such as those employed in the stripping of negatives for offset printing. These sheets have a slight roughness on one side which allows the scorer to draw a series of circles around the right answer positions in drawing ink. When this is used as an overlay, it is possible to see the entire candidate answer sheet through the plastic and thus double check that credit is not being given to multiple-marked answers. In addition, this method makes it easier to differentiate between an intended response and a poor erasure. However, some colleagues to not agree that this improves the efficiency of scoring since the eye cannot pick out the right answers as readily as with the usual punched scoring stencil, and the two methods have yet to be compared empirically.

The second modification is to score the second side of a two-sided answer sheet first. Count the number of correct answers, turn the sheet over to the first side and write the number correct at the top of the sheet. Then in the second run through, score the front side, write the number correct underneath the other score, add the two scores, and circle the total. This eliminates much of the time and error involved in transferring the mark from the second side to the first.

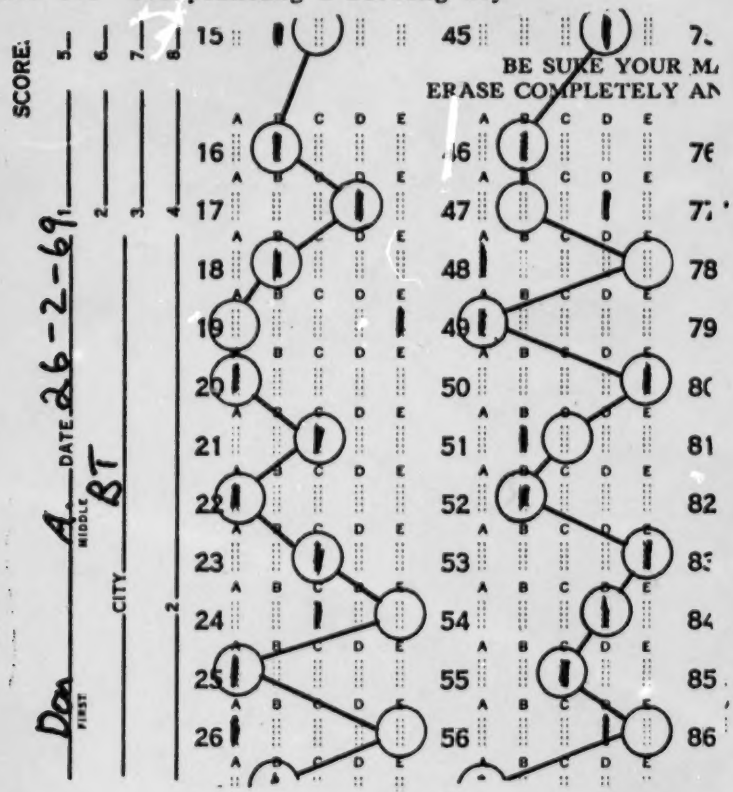
A much simpler but also much more expensive scoring method involves using self-scoring answer sheets. These are typically two sheets of paper bound together with a carbon paper interleaved. The candidate marks the top sheet, and the carbon transfers his marks to the second sheet where a scoring grid has already been printed. The marks appearing in the preprinted circles are then counted.

Another, more sophisticated, version has the correct answers marked invisibly with a special solution. As the candidate indicates his answer with a special pen, the mark he makes changes color if the answer is correct.

Because both of these are more expensive than the previously described methods, they are not as widely used, but this is only one reason. A second reason is that most testing offices would be extremely uneasy and fearful that some supervisors were having a single candidate find the correct answers and pass this information on to others. When the only answer key resides in their safes, officials worry much less about this problem.

A variation of the self-scoring approach is for the testing staff to prescan all answer sheets and then to overprint a set of circles around the correct answers, using either green or brown ink for less eyestrain. A mimeograph can be used if care is taken to ensure correct registration. (See Figure 8.1 below.)

Figure 8.1 Overprinting a scoring key



But the fastest and most accurate scoring is that carried out by machine. In addition, as the number of tests to be scored increases, the relative cost of using the computer to do it decreases. In the past few years, a variety of machines for reading answer documents optically have been introduced into testing offices. Two of the largest manufacturers,

IBM and ICL, have page readers which can handle the volumes of answer sheets in most large-scale examinations. The IBM 1230, for example, can mark about 1,000 sheets per hour, make corrections for guessing, print the subscores and total scores on the paper, and punch the results into Hollerith cards if needed. The ICL UDT is faster, but must be connected to a computer. Other manufacturers make far superior equipment, but do not have as many world-wide facilities for maintenance.

In recent years many offices have begun to prefer answer cards, since some of the available equipment for scoring these is small and simple, the cards themselves are more resistant to the wear and tear of shipping and usage, and the processing speed is much greater. For instance, the Hewlett-Packard 2761A optical card reader will handle 200 cards per minute, which means a processing rate of 12,000 per hour; the cost of a complete small data processing system with optical card reader, computer, and printer, is less than that of the IBM 1230/534. For extremely large-scale examinations, such as the West African secondary school entrance examinations which about 100,000 candidates sit per year just in Nigeria, and each with four papers, making a total of 400,000 papers needing to be scored, the older methods simply cannot compete with the card reader.

And with the card reader connected directly to a computer, an office can program the equipment to perform comprehensive edit checks as the answer cards are being processed. This allows the staff to make corrections immediately, and keeps the quality control at a high level throughout the various steps.

To bypass the maintenance problem, the Malawi office uses equipment with an average time between failures of over 2,000 hours, and ships the reader (28 pounds shipping weight) outside the country for regular servicing.

Machine scoring simplifies the matter of correction for guessing, or differential weighting of item responses. Any formula can be applied and marks easily adjusted. In spite of the ease of applying a formula, the office rarely does. Although some objective testers advocate correction for guessing as a counterargument to the traditionalists' claim that candidates can pass objective examinations by guessing, the STO still does not use it, and the reasons are simple. The first is that so far it has not made any difference to the order of merit in West Africa. The second is that it complicates the procedures, making them a time-consuming, error-producing

process when the machine scoring is checked by hand.

The West African office stumbled onto the fact that correction for guessing made no difference in a typical STO manner. The directions for one of the larger examinations stressed that correction would be made, and the candidates were told: "You will be penalized heavily for guessing. DO NOT GUESS!" The instructions for making a correction for guessing, however, were not sent to the data processing section until after the results were already printed. Then the quandary: should the office redo the entire examination and insist that over 500 secondary schools begin late, or should they accept the results and take a chance that they might be caught in their deception?

To assess the extent of difference the error would make in the order of merit, they took a sample of 5,000 papers and rescored them with correction for guessing. The correlation between the uncorrected and corrected scores was .997; i.e., the correction made no difference. So they decided to issue the results as they were. When the author asked one of his Nigerian colleagues how he would explain this phenomenon, he laughed and said, "When you tell a candidate not to guess, he ignores you. He knows he has only a tiny chance of being accepted for secondary school and that he just might guess lucky, so he guesses. EVERYBODY guesses. Everybody plays the pools, everybody takes a chance, everybody guesses."

Two particularly important lessons were learned from this experience: Assumptions made in Western cultures may often be untenable when transplanted to non-Western soils, and the directions for processing the examination results must be made clear and explicit before the processing begins.

Some sort of repeat scoring is imperative no matter what scoring method is used. With machine scoring, a set of test papers is sent through the machine twice and the two sets of scores compared; then a sample of this set is hand scored and the marks compared with both the machine-printed ones. This procedure is repeated constantly during the machine scoring. One operator does nothing but select and recheck samples at regular intervals.

If an optical mark reader is used with a computer, the computer can be programmed so that it stops every few minutes. At this point the operator must rescore the last examination paper processed. If the computer and the operator agree, the computer is allowed to continue; otherwise the discrepancy must be corrected. Repeat scoring is part of the process called quality control--vital to the entire office's working.

If the papers are being hand scored, the first set of scores should be completely hidden while a second scorer rescores the same papers. The two sets of scores, arrived at independently by two different scorers, are than compared. In this way the accuracy of the scoring method can be determined. If such a check discloses that a scorer is in error, *all* of his papers must be rescored.

Despite all efforts to keep it down, the error rate for hand scoring is insomnia-producing. Without confessing the actual rates obtained in practice, an idea of the magnitude can be given. On one examination every paper was machine scored. If the number of poor erasures on an answer card exceeded one, that particular card was hand scored since the equipment being used could not differentiate between the proper mark and the improper erasure.

However, a formula was used to estimate the number of improperly erased questions the candidate would have had correct *if* he had erased properly. Senior officers then carefully re-marked the cards twice. When the clerks' hand scored marks and the machine estimates were compared with the officers' scores, the machine estimates were found to be significantly closer to the scores the officers obtained than were the hand scored marks. This may not be convincing to anyone who realizes that there is little relationship between intelligence and accuracy of scoring.

No matter how accurately the scoring is done, there are invariably inquiries about a particular paper. If the answer sheets are classified by test center, with all the papers from a particular center kept together, specific papers can be located much easier than if papers from all centers are stored together. Although this sounds like a simple point, at one office hundreds of man-hours had to be spent in a fruitless search for 200 papers among a quarter of a million before the procedure was changed.

An important part of the processing is checking the quality of the data at each step. Although the general problem of quality control is discussed in greater detail later (and in section 9), a few checks are introduced here. Repeating the scoring of papers has already been mentioned, as has the general edit of all the scores, which is a simple task if the computer is used for processing. The minimum editing which must take place is inspection of all the scores to see that they are all within range. That is, if the maximum possible score for test A is 60, no candidate can have a score above that value, or below 0. Out-of-range scores do occur

with help from the clerical staff and occasionally malfunctioning machines. There should also be some consistency to the marks on all the papers which a candidate takes. If he has an extremely high score on one paper, his other scores should be reasonably high as well. Cases of gross inconsistency may point to defects in the examination, or in the scoring system.

A recent case may illustrate this. A visual scan of scores for candidates on four papers showed that about ten individuals at one test center had inconsistent results. Although their marks in papers A and C were high, their marks in paper B were low. Upon investigation it was found that the scores from the back side of paper B had never been added to those on the front side .

But absentees present the major problem at this stage of processing, and any errors in identifying them may lead to acute embarrassment. How does one explain away the high pass mark which a long-dead candidate receives? The only reasonably foolproof method of avoiding such problems is to handle absentees with a special code. The list of those absent and present at the examination hall is used to assign "A" marks to those absent. Then every single candidate *must* have a mark for every paper: either an "A" or some legitimate score. A blank is treated as an error and the lost paper must be found. And if an answer sheet for an absent candidate turns up, something is decidedly wrong somewhere.

When more than 13,000 candidates were apparently absent out of total of 51,000 in West Africa, the senior staff were naturally suspicious and did some checking. Someone had forgotten to check the absentee lists and had merely assumed that a missing score meant that the person was absent. Introduction of the absentee marks of "A" left only 11,000 unresolved cases which were largely resolved when the shipment of answer sheets was found at the wrong freight depot. After scoring these, only nine unknown cases were left, the closest the office had ever come to perfection, and that after being so near disaster--the results were being released the following day!

Since each test is marked separately, at some stage all of a candidate's marks must be brought together. This merging (often termed "marrying") of the scores from the various papers is fraught with possibilities of error. An entry file, or master list of candidates, is matched with the lists of results and the scores on each paper transferred to the corresponding entry. If the matching and transferring are done by hand, a candidate may be missed

and all the subsequent marks given to the wrong individuals. If this is done by computer, the programmer may not have included checks of the student number.

This latter error was made at one unfortunate testing office, and when coupled with their lackadaisical handling of absentees, almost permanently ruined their reputation. Some candidates received the marks of others, obtaining school certificates sometimes with passes in as many as 22 subjects. Since candidates were allowed to take between six and ten subjects, the office personnel should have spotted this error easily, long before the results were released. But they did not. Imagine the delight of the many students who found they had passed with distinction subjects which they had not sat, and the dismay of others who were considered absent for papers they had completed. The ensuing chaos compelled the testing office to redo the processing completely (and to sack the irresponsible officer). The results were announced almost two years late.

The combining of marks from objective examinations and essays or practicals proves no more difficult than combining only objective marks. The major trouble is the comparatively greater length of time needed for the marking of the essays and practical examinations compared to that for objective tests.

In some places, although the essays are given, most are not marked. For instance, the papers for the candidates on the margin between pass and fail may be the only ones read. Or the papers of the top ten per cent may be sent to the headmasters of the secondary schools to which the candidates applied. The headmasters then have these essays marked, and the combination of objective and essay marks is used to determine who will be asked to come in for interviews.

Essays can also be used to detect impersonation in the examination. If he is suspicious of a candidate, the headmaster may ask him during the interview to write a short paragraph on some subject. He can then compare the candidate's handwriting on the paragraph with that on the essay. Often it is different; the candidate may have had his elder and more clever brother take the examination for him. Where this method of detecting impersonation is not possible, the use of pictures as described on page 2-2 is preferred.

The methods of marking traditional essay examinations are patterned after the methods developed in England, and are roughly as follows:

1. Written instructions are given to all the markers and these are supported by discussion and further briefing before the actual marking.

2. Papers are marked by individuals under the general supervision of a team leader or chief examiner.

3. Borderline or otherwise doubtful papers are re-marked under the chief examiner's direction.

Additional information on this subject is available through most of the UK examining boards.

Often when the separate marks are combined for each candidate, the testing office or moderating committee prefers to inspect the resultant distributions before deciding on cutoff points. If a computer is being used, the calculating and printing of frequency distributions is a simple matter. After inspecting these, if the committee has some a priori idea of an appropriate passing percentage, they can easily decide on a pass-fail cutoff point. A computer-printed frequency distribution is reproduced below (Figure 8.2)

Because of its objectivity and efficiency, the frequency distribution method for deciding on pass-fail c toff points is preferable to the usual methods of inspecting candidates' scores in order of merit. In cases where the committees looked at the composite results listing to decide on the cutoff point in the past, the names and home villages of the candidates were shown, and the members may well have been unconsciously affected by these in their choice. Even if the identifying information is removed, the frequency distribution method is preferable because it gives the same information in less time. In one case, the STO decreased the meeting time for a committee of 20 members from three days to one and a half hours when they switched from order-of-merit listings to frequency distributions.

If transformations of the data are required, they can be done at this stage. Various committees and offices have their own ideas of the relative weights each paper should carry in the composite total, and transformation of the marks to some arbitrary mean and standard deviation simplifies the process. (If equal intercorrelations of all the papers are assumed, this matter is vastly simplified.) Then each transformed mark is multiplied or divided as desired, and the total is calculated for each candidate.

As an example, assume that the marks on the ABC examination were all changed to an arbitrary mean of 500 and standard deviation of 100. If the committee or testing office desire to give the C examinations a double weight, the new scores on C can be doubled before adding A and B. If Mr. Mercury

Figure 8.2 A computer-printed frequency distribution.

SCORE	NUMBER	PERCENT	CUM N	CUM %
48	2	.02	2	.02
47	3	.03	5	.05
46	3	.03	8	.09
45	6	.06	14	.15
44	6	.06	20	.22
43	0	.00	20	.22
42	0	.00	20	.22
41	11	.12	31	.35
40	8	.09	39	.44
39	9	.10	48	.54
38	11	.12	59	.66
37	15	.16	74	.83
36	19	.21	93	1.05
35	22	.24	115	1.30
34	19	.21	134	1.51
33	34	.38	168	1.90
32	42	.47	210	2.37
31	62	.70	272	3.07
30	51	.57	323	3.65
29	83	.93	406	4.59
28	102	1.15	508	5.75
27	112	1.26	620	7.01
26	133	1.50	753	8.52
25	166	1.87	919	10.40
24	173	1.95	1092	12.36
23	205	2.32	1297	14.68
22	225	2.54	1522	17.23
21	248	2.80	1770	20.04
20	276	3.12	2046	23.16
19	328	3.71	2374	26.87
18	326	3.69	2700	30.57
17	384	4.34	3084	34.91
16	386	4.37	3470	39.28
15	380	4.30	3850	43.59
14	419	4.74	4269	48.33
13	480	5.43	4749	53.77
12	427	4.83	5176	58.60
11	408	4.61	5584	63.22
10	429	4.85	6013	68.08
9	404	4.57	6417	72.65
8	344	3.89	6761	76.55
7	346	3.91	7107	80.46
6	261	2.95	7368	83.42
5	219	2.47	7587	85.90
4	181	2.04	7768	87.95
3	153	1.73	7921	89.68
2	74	.83	7995	90.52
1	42	.47	8037	90.99
0	795	9.00	8832	100.00
7	BLANK			

received transformed scores of 530, 465, and 640 on the A, B, and C, papers respectively, his composite (or aggregate) score is 530 + 465 + (2 X 640) = 2,275.

The advantages of transforming are many: the transformations aid in interpretation, may point out instances of test leakage, and often are more similar to committee members' conceptions of the results than are the raw scores.

That interpretation is easier can be seen from the Headmasters' booklet, page 1-14. Once a system has been decided upon, to continue using the same mean and standard deviation simplifies the communication of results to the users. A headmaster can say to himself, "It looks as if those applying to my school this year are not the best students from among all the candidates as they were last year."

Another use of transformed scores is to detect examination leakage or administrative misconduct. When the more difficult items are chosen from the pool of questions and assembled into a test, most candidates are unable to guess what a good score would be. If they clandestinely obtain a copy of the examination, they will likely study every question and as as result do very well--in fact, too well.

In practice, the computer is programmed to print out a list of all candidates and their schools who receive scores above 800 (see page 1-14). If the list shows any two candidates living within a 50 mile radius, they are investigated. One time twelve candidates from a single school scored in excess of 800; two of them were above 1,000. When the Criminal Investigation Division finished checking the facts concerning this highly unusual event, three teachers were in jail.

In some places, the moderating committee prefer transformatons which resemble percentage marks, which are more familiar to them. The use of a mean of 50 and standard deviation of 16 satisfied one group since they felt that about 20 per cent should fail each paper, and 36 had been the passing mark in the past.

When the decisions concerning cutoff points and the relative weighting of the papers have been made and the various scores added together, the candidates can be sorted into some order of merit and listed. (See Figure 8.3) Political considerations take precedence in most countries, so the order of merit could be by district, by region, by type of school, or whatever subdivision satisfies the local conditions. Again, if the computer is used,

NIN ED 1970 PAGE 3

CENTRE 104 GOOD SHEPHERD

CLASS	NAME	ID	SEX	AGE	ENGL(200)	MATH(150)	SESU(100)	GENL(50)	TOTL(500)
					PERCENTAGES				
3 RD	MOKOLATSIE MOKHELE ELISHA	17	M	19	43.0	47.3	51.0	20.0	.4
3 RD	MOLELEKOA TSIETSI JOHN	18	M	18	51.0	64.0	38.0	34.0	.5
2 ND	MOLISE NTHABISENG AMELIA	19	F	15	51.0	54.0	65.0	38.0	.5
3 RD	MOLISE NKUNE BERNADETTA	20	F	17	50.0	34.0	54.0	34.0	.4
3 RD	MOSELANE KOTOPE PASCALINA	21	F	16	43.5	38.6	67.0	20.0	.4
3 RD	MONONGOAHA MAETI AUGUSTINA	22	F	17	50.0	40.0	66.0	36.0	.4
3 RD	MOTHEPU 'MASEPONTI VALERIA	23	F	17	45.0	43.3	51.0	30.0	.4
3 RD	MPATS'EHLA 'MATSABA MARY	24	F	16	42.5	34.6	62.0	22.0	.4
3 RD	MPHALE NTS'IOOA CLAURINA	25	F	13	52.5	46.6	45.0	22.0	.4
3 RD	NKOLO TLALENG BERNADETTA	26	F	15	48.0	52.6	63.0	34.0	.5
3 RD	NTEPE JOBO FRANCIS	27	M	17	42.5	28.6	54.0	30.0	.3
2 ND	NTEPE PALESA MABEL	28	F	13	60.0	44.0	68.0	48.0	.5
3 RD	NTHUELE MATHEKO JOSEPHINA	29	F	18	55.0	24.0	54.0	58.0	.4
1 ST	NTJANTJA THUSO PASCALIS	30	M	16	64.0	90.0	62.0	34.0	.6
2 ND	PENENE KHOBELO CECILIA	31	F	19	54.5	54.6	62.0	24.0	.5
3 RD	ATE MAMPEMI FRANCISCA	32	F	17	56.0	38.6	65.0	26.0	.4
2 ND	SOTS'O 'MAMOCHINI ANNA	33	F	17	59.5	43.3	65.0	2.0	.5
****	LEBOEA LISEBO AUGUSTINA	34	F	17	38.5	27.3	40.0	28.0	.3
3 RD	RANYALI MAMOEPA ALPHONSINA	35	F	16	40.0	40.0	50.0	38.0	.4
3 RD	RAPHOTO MPOEEA MARCELLINA	36	F	16	50.0	40.6	65.0	22.0	.4
3 RD	SEBITSA KHOMANANG CAMILLA	37	F	17	49.0	40.6	58.0	28.0	.4
****	SEKELEOANE MAMOKETE JUSTINA	38	F	16	0.	0.	0.	0.	0.
3 RD	SEMPE 'MANTS'EBO AMELIA	39	F	18	47.5	37.3	61.0	16.0	.4
3 RD	SENYANE LITLALLO EXINIA	40	F	18	40.5	58.0	44.0	26.0	.4
****	TEMA AUGUSTINA TSOTANG	41	F	16	0.	0.	0.	0.	0.
3 RD	THETHA 'NOKOANA CECILIA	42	F	18	51.0	52.6	57.0	38.0	.5
3 RD	TSOAELI MOOKHO JUSTINA	43	F	20	50.0	54.6	61.0	32.0	.5

1 1-STS. 7 2-NDS. 29 3-RDS. 6 FAILS. 43 TOTAL 45.2 39.3 53.5 28.5 43.4

CENTRE 105 HERMON

CLASS	NAME	ID	SEX	AGE	ENGL(200)	MATH(150)	SESU(100)	GENL(50)	TOTL(500)
					PERCENTAGES				
****	BOHLOKO ELLEN	1	F	17	50.5	11.3	56.0	26.0	.3
2 ND	MELELE SIMON	2	M	20	54.0	44.6	58.0	66.0	.5
3 RD	MLAISWAYO NUMASONTO	3	F	15	54.0	34.6	50.0	46.0	.4
3 RD	KEKETSI KEFUOE	4	F	16	41.0	28.6	55.0	46.0	.4
****	KHAKHANE MARIA	5	F	16	0.	0.	0.	0.	0.
****	KHAKHANE MOIPONE	6	F	16	0.	0.	0.	0.	0.
		7	F	16	[illegible]	[illegible]	[illegible]	[illegible]	.5

8-13

Figure 8.3 Results listing

the actual printout of results can be used to make offset plates so that no hand typing need be done. In previous years the results in Botswana had been typed on mimeograph masters--over 6,000 names and grades each year, an enormous task which took a battery of typists several weeks to complete. In 1970, the computer did it in 30 minutes. An exact copy of a page of results is shown in Figure 8.4 below. The listing is by school, before any sorting took place. Personally, the author objects to an order of merit within schools for such a listing since it is unnecessary for a candidate to know that he was the lowest of the failures, or the highest of the passes. In fact, no marks were shown for any candidate, so only the Ministry of Education can make comparisons among the schools in more detail than the number of A's, B's, etc. (The Ministry received a more complete printout similar to Figure 8.3; the printout in Figure 8.4 was public information.)

The printing of certificates can also be done by computer, if one is available. One advantage of using the computer is that its distinct print is extremely difficult for forgers to reproduce. Additionally, it decreases the number of typographical errors, and speeds up the process of issuing the certificates immensely.

Most of the certificates were printed on special security paper. Any attempt at erasure, either using a rubber or a razor blade, would expose a sublayer of paper on which was printed "invalid." However, the paper was very expensive and had to be obtained from Europe. To keep the costs down as much as possible, a large supply was ordered and the date was inserted by the computer; thus one order would last for many years. As an additional security measure, the certificates were numbered so that every one could be accounted for, and no STO signatures were affixed until after the candidate's names were printed. Then a signature block was taken from the safe, and the official signature overprinted with a letterpress on all the certificates. Thus someone would have to obtain a blank certificate and falsify the number register, print a name on it using a type similar to the computer's, and then overprint the official signature in order to forge a valid-looking certificate. Another advantage of this method is that the signature block could be changed whenever a new official took office without disposing of all the unused certificates.

In the countries with even smaller budgets, the certificates were printed on ordinary paper. To increase the security, a type of paper was

Figure 8.4 Public results listing

BAKWENA NATIONAL

GRADE	CANDIDATE
	CANDIDATE
C	BABUTSI MAY
C	BILE RAMOLEFE
C	GAITSIWE KOOLEBILE
C	JOSEFA BONTLE
C	KABOYAKGOSI SOPHIA
D	KAGISO LEGAKABE
C	KEDUMETSE ONTIRETSE
D	KEJELEPULA POMATSHEGO
B	KESEBONYE MOLEFE
B	KGOSIDIJO NEO
A	KOKORWE BATSHEDI
C	LESELAMUSE KEITUMETSE
C	MANGATE KEMAKGAMETSE
B	MASIELE ONNEETSE
C	MASISI LERATO
D	MOGONE LATLHANANG
D	MOITOI ONALENNA
C	MOJALEMUTHO GLADNESS
B	MOKGOSI AMELIA
B	MOLATLHEGI MOTSHEGETSI
B	MORENANG GODFREY
B	MORNAENG JOSEPH
B	MORNAMANG KEITUMETSE
B	MUSARWA GRESI
C	MUTHUPI SISTER
B	MUTLHABI MODIRWA
B	MUTSHADIBE KELEBONYE
C	MUTSHEGWA GABONA
C	MUTSWAKUMO GOABELETSWE
B	MUTSWEDI LEVY
B	NTWAETSILE TUOUETSO
D	RAMAGOBO DITSRO
C	RAMPUTSWA VIOLET
B	SEBETSO ALFRED
B	SERUNYA SARAH
D	TIRELO DINEO
C	TIRELO MERCY
B	TLADI SELEMELA
C	TSUAMENG MOLIFA

CANDIDATES= 40

BORITSWANA

GRADE	CANDIDATE
	CANDIDATE
C	CHABAESELE ITUMELENG
B	CHEPETE IKOTLHAENG
B	DIFELE RAESANG
B	GABOSEGWE DINAH
B	GOKGAKILE OLHAPILE
B	GOSGANNANG RAESANG
B	GOPOLANG SIUKILE
B	KABLAY JACOB
C	KEITAMTSENG SEKGOPI

GRADE	CANDIDATE
	CANDIDATE
D	RASEDIE ELSIE
A	RASEDIE LEEPO
B	RATLADI VIRGINIA
C	SEGWATI AMULE
B	SEITEI PATRICIA
B	SELEKE DOREEM
D	SELEKE MAGDELENE
B	SELEMOGWE BUITUMELO
B	SIPA MOIPONE
D	TABANE PHOKWANE
B	TMOKWANE SETSHEDI
C	TLADI ANNAH
C	TLHANKANE NKIDI
B	TLHASE MPOPO
B	TSATSI NICHOLAS
A	TSATSI SOLOMON
C	TSELE SETLHALEFENG
B	TSHEKO MUSADINYANA
B	TSHEOLE SOPHIE
D	TSIBELE DURAH

CANDIDATES=100

SIKWANE

GRADE	CANDIDATE
	CANDIDATE
B	DINTWE MALENYORA
B	GABORONE VICTOR
C	JAUTSE JOYCE
C	KGAOGI MARIA
C	KGWADI LAZARUS
C	LEKORWE SARAH
C	LETSEBE MOLEFE
A	MFETANE KGOMOTSO
C	MODIBE MARGARET
C	MUDISE JOCOBETH
B	MOLEFE KELETLHORILE
D	MORUBE THOMAS
B	MUTSILANYANE LEVI
C	MPETE ISAAC
C	NTWAGAE RUTH
B	PHETLHE ELIZABETH
C	PHIRI NAOMI
C	PHIRI ORPHA
B	PILANE ICABOTH
B	RATAU BOYCE
C	SEABELO MOAKI
C	SEABELO SULUMON
C	SEEMISE BERTHA
D	TMEBE GRACE
D	TMEBE ICABOTH
C	TLHAGWANE CATHRINE
C	TLHOME CATHRINE

CANDIDATES= 27

GRADE	CANDIDATE
	CANDIDATE
A	MOTLALEKGOSI DICK
B	MOTSOMI PHINIAS
C	NKWE DIKELEDI
C	NTSHOLE MORARE
D	NTSHOLE MPOLOKENG
C	NTSHOLE REBECCA
B	PITSE ONICAH
C	SEBOLAI DINAH
C	SENTSHO JANE
B	SETLHABI LAZARUS
B	THAMAGE JOCOBETH
C	THAMAGE SELLO
C	THAMAGE SELLWE
C	THAMAGE SETH
B	THAMAGE STEPHEN
D	THIBEDI JOHN
B	TLADI KGOMOTSO

CANDIDATES= 34

SEGALE

GRADE	CANDIDATE
	CANDIDATE
A	BAKWENA ALICE
A	BAKWENA MERCY
B	CHELANE MELEKO
C	CHUMA MERCY
B	DITHARE NTSHADI
A	DITSHWENE MORENWE
B	DITSHWENE SERGETHELO
B	GARE MOSIDI
B	GARE RADISE
B	KGARI VINCENT
C	KOPONG BAATSEBA
A	KOPONG NTSWANE
D	KOPONG PETER
C	KUTLAPYE TLADI
B	LEAGAJANG NTWARI
B	LEBURU SEISABOGADI
A	LESEJANE SENEMA
A	LETSHOLO MMATSELA
B	LETSWITI MUDIANE
C	LINCHWE TSELANE
B	MAGOLE MARGARET
A	MALEBYE VIRGINIA
A	MARIBE MOLEFE
A	MARIBE ORATILE
B	MASOKO STANLEY
C	MATLAPENG KEDISITSE
B	MATLHAGA KGOSEDI
A	MATLHAGA MASEGO
B	MATSHEDISO MUSIDI
C	MATSHEDISO WILLIAM
D	MMAL- MUTLALETSI
B	MUDISE MOLATLHEGI
B	MOEKETSI DAVID

GRADE	CANDIDATE
	CANDIDATE
B	MOEKETSI LESEGO
B	MOEKETSI MMAMEA
C	MOENG PUNE
B	MOGALE BOTHOBOILE
C	MOGAPAESI GUITSEMODIMO
C	MOGOTLHWANE TSWETLHE
B	MOGUTSI BARATI
C	MOGOTSI VELILE
A	MOJI MMATLADI
C	MOKONE DITSILE
B	MUKONE WILMA
B	MUKOTEDI SEGALE
C	MOLATLHEGI KGOMOTSO
B	MOLEFE KGOMOTSO
C	MOLEFE MMATSELA
A	MOLEFE MORENWE
C	MOLEFE MUTLATSI
A	MOLEFE SELEBI
A	MOLEFE TENNYSON
D	MOLENI JACOB
D	MOMGALE MAGAPE
B	MONONYANE SEKOLWANE
A	MORUAKGOMO MOTSEI
C	MOSENYI MKOMENG
C	MUSIMANENAPULA GAONGALELWE
B	MUSIMANENAPULA TUELO
C	MUSWEU MATLHODI
B	MOSWEU MMAKONE
B	MUTLHABANE DUNIS
A	MOTLHALE MODIANE
B	MOTLOGELWE MMAPULA
A	PALAI GRACE
B	PALAI MASILU
B	PALAI STEPHEN
B	PMALLAGAE WILLIAM
B	PMATSWANE MMATSHIDI
A	PHETO ESTHER
B	PHETO PHETO
A	PHIRI JANE
A	PHIRI LYDIA
B	PHIRI MOTHULWE
C	PHIRI TEICHLER
C	PILANE MALEFSANE
B	PILANE TMOBOLE
C	POONYANE ABEL
B	POONYANE JOSEPHINE P.
C	RAKERENG DIKULO
B	RAKERENG TSIEPE
B	RALETEBA DIKELEDI
B	RALETEBA OTSWILE
C	RAMALEPA KAILANE
A	RAMEDUPE MOTLAPELE
B	RAMODIANE MMAKOBO
B	RAPALAI MMANYLO
C	RAPALAI MORMUTE

used which was comparatively heavy and also unavailable in the country. Around the border were added some embellishments, and the entire certificate printed in an off-color green ink. The printing of the candidate's name, when possible, was done using a matrix printer so that each letter was composed of a series of small dots. The combination, it was hoped, would prove difficult to duplicate. Figure 8.5 is a black and white copy of a completed certificate; the cost of each one, including printing of the candidate information, was approximately 1.2 cents.

Figure 8.5 A Primary School Leaving Certificate

EDUCATION DEPARTMENT
STANDARD 7 CERTIFICATE

30 November 1970

Republic of Botswana

This is to certify that:

BALESENG DIMAKATSO of BOBONONG

School has written the Standard 7 Examination and has received a Grade of:

C(AVERAGE) Number: 1002.

D.W.Kgaleb

Chief Education Officer

Candidates who lose their certificates and request another receive a new one clearly marked "duplicate." The charging of a nominal fee for the issuing of a duplicate certificate is favored since it takes so much staff time and paperwork to prepare a new one that a fee may discourage carelessness. If the design of the certificate changes from year to year, an excess of certificates should be printed each year to allow for those lost.

Publication of the results should be as widespread as possible, and as soon as practicable after the examination. The longer the delay, the greater the tendency for impatient candidates to telephone, write, or even stand around the STO in hopes of finding out the results. Many countries publish the names of successful candidates in the official gazette, or in the newspaper. If the STO were so well organized that the results could be released on a predetermined date, its fame would spread throughout the world.

Some offices which test a large number of private (non-school) candidates are troubled by the problems of disseminating results quickly. One possibility is for the STO to include blank envelopes in the examination shipment; the private candidates address these to themselves at the time they take the test. The Schools Examinations Council in Lesotho uses this method since the majority of private candidates change their addresses between registration and final results reporting.

But the attempt to get the results out promptly should not preclude the careful check of all steps in the process; quality control must continue to the very end. The last and final check should be a careful perusal of the results list by a senior official of the STO. If this step were followed consistently, a great many offices would be in a lot less trouble. An example from the U.K. might be more convincing than a dozen from elsewhere: one large testing office issued thousands of identical O-level results of 9's in mathematics. Almost every single candidate had a nine, and a simple visual check would have stirred an official to ask if the standards had dropped drastically or if the results were seriously in error (as of course they were). Such checks would have spotted the majority of the errors referred to in earlier sections of this handbook.

A number of simple quality control techniques can be used throughout all the processing. One highly effective one is to insert known "defectives" into the process. These can be names of ficticious candidates with no corresponding answer sheets, or answer sheets with no corresponding entries. Candidates with too high scores, or with clearly inconsistent scores can be included. A similar technique is to remove a few names or papers. In all of these situations, if the appropriate section of the STO does not locate the fault, beware.

Some of the steps, especially the manual ones, can be checked by a simple repeat of the operation by another person. Care should be exercised that there is no communication of the first person's results to the second person, however. Repetition of the scoring, transforming, or merging can be helpful in identifying the weak spots in the total system.

Repetition of the scoring at the request of the candidate sometimes occurs. The procedure which one office uses is reproduced here:

(a) A candidate who has failed at the examination in one or more subjects may have his script in such subjects re-examined on payment of a fee of $4.20

per subject, provided that application for such re-examination is made within one month of the date of publication of the results. The fee for re-examination in any subject will be refunded if the candidate is allowed to pass in such subject as a result of the re-examination.

(b) A candidate who has passed at an examination may have his script in one or more subjects re-examined on payment of a fee of $4.20 per subject, provided that application for such re-examination is made within one month of the publication of the results and the STO approves of the purpose for which the re-examination is requested. In such cases the fee for re-examination is not refunded.

But good supervision is probably the key element in ensuring the highest accuracy and speed in processing. Since this is such an important task, some STO's try to have some of the senior staff assist in supervising the regular processing staff. They also set up the various steps so that they can identify the particular person who made the mistake. In one case where the personnel were asked to sign the papers after marking them, the error rate dropped from six to less than one-half per cent.

Unfortunately, no matter how much care is taken, some errors show up when the summary statistics are determined for the entire examination. The mean, standard deviation, number of applicants, number of passes and failures, and intercorrelations of the papers provide some additional clues to the accuracy of the processing as well as give information useful in planning for subsequent years.

For some examinations, such as school certificates, information on the numbers of individuals receiving various combinations of passes is helpful to both the ministries of labor and education in their planning for further training. The exact nature of the statistics reported vary considerably from country to country, but the ones mentioned above can be considered a minimum.

If the results are accurate, they are worth keeping. Retention of the data makes possible the assessment of validity; e.g., comparing secondary school selection results with school certificate results obtained four or five years later. If the data are placed on magnetic tape, retention for 50 years is a simple matter. Some retention is always necessary to provide the means for checking requests of candidates for duplicate certificates, and occasional inquiries from prospective employers about the legality of

a particular certificate.

Essentially, when the statistics are calculated and the results lists placed in storage, the cycle is complete. In many testing offices, advertisements for the next examination have already appeared in the newspapers, applicants have started registering, the pretesting results have been analyzed, and preparations have started for another round of scoring, processing, and reporting.

OVERALL SECURITY

A clever candidate once wrote to a friend in another country where the examination he was preparing to take was being printed. The obliging friend investigated the contents of the printers' waste bin for a few days and was rewarded with multiple copies of the examination to send back. This incident, with slight variations, has occurred in at least three African countries in the past five years, and is, of course, just one example of the hundreds which occur. In fact, the number, variety, and ingenuity of plots devised to outwit examination boards have always been so astonishing as to seem nearly mythical, and they continue to increase--in number, variety, and ingenuity. For this reason, a sense of the need for security must pervade every crevice of a State Testing Office or it is doomed to failure. When the first intimations of mistrust are cast upon the office's ability to maintain the confidentiality of its examinations up to the moment of administration, its reputation is imperiled. Once it is suspected that one or more candidates obtained knowledge of the test content, and "got away with it," the examination results will be discredited for and by everyone.

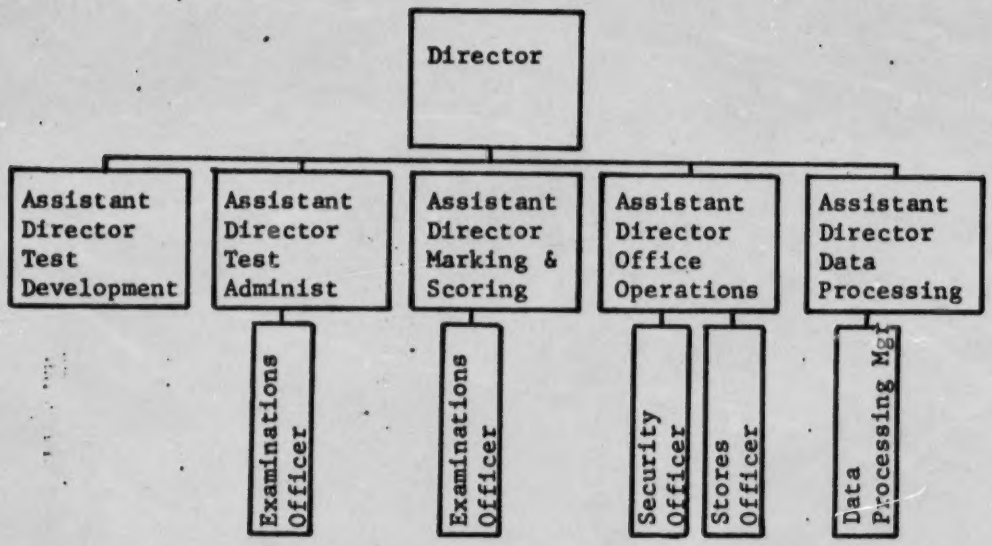
If the testing office has a set of standard operating procedures for dealing with security materials, it can avert a great many disasters, and quickly trace any errors which occur. Procedures might reasonably be expected to differ from country to country, depending on the physical accommodations at the testing office, the acceptable modes of transportation, and the modus operandi of unscrupulous candidates. In this discussion, a set of specimen regulations are offered which may require modification depending upon such variables. As a background for the regulations, imaginary plans of a typical testing office are presented, some possible staff positions are specified, the security areas are shown on a floor plan, and a few assumptions are made about availability of equipment. Frequent reference to this fanciful organization should help the reader interpret the sample materials (see Figures 9.1 and 9.2). For added clarity, the material is subdivided and subheaded.

I. Security Areas.

The areas of maximum security consist of a number of specially constructed rooms and conventional office safes (or heavy lockable file cabinets).

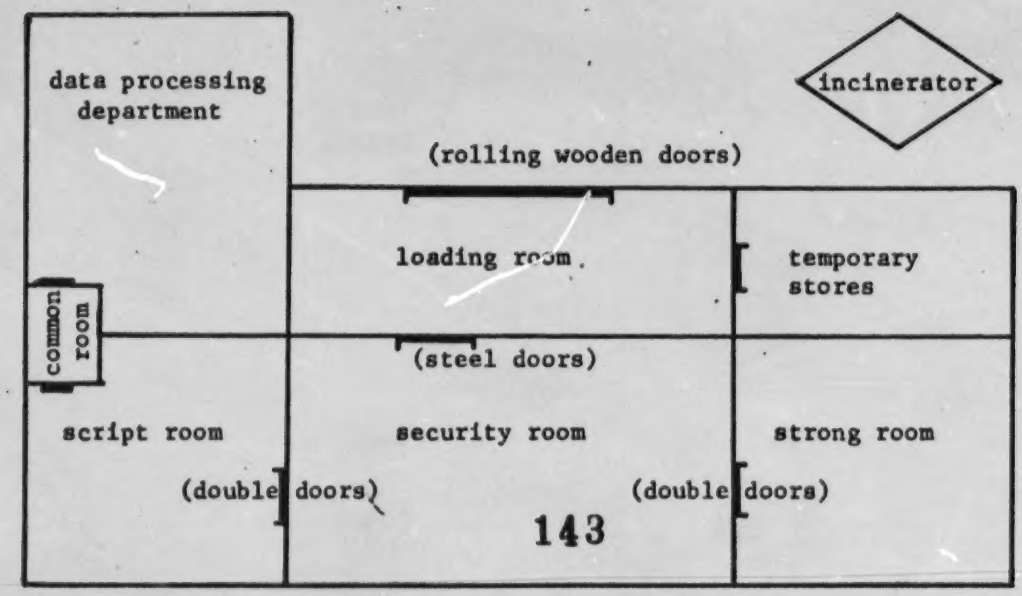
A. Rooms. The rooms are: (1) a script room used mainly for the storage and transfer of completed examination papers and results; (2) a security room used mainly for storage and transfer of question papers; and (3) a strong room

Figure 9.1 Office organization.



Only a small portion of the entire organization is shown.

Figure 9.2 Security area floor plan.



used only for storage of unopened containers of security materials. In addition, a temporary stores room adjoining the loading area is used for storage of minimum security materials such as unused answer sheets, apparatus for practical examinations, etc.

The security room is divided into two sections, an outer area which is used for the offloading and loading of vehicles, and an inner room which is used as a security room. The main doors to the loading area, when locked, provide complete protection for any material that may be temporarily stored there. These doors can only be opened and locked from inside.

There is no direct connection between the security room and the loading area, although materials for storage are passed into the security room through an aperture in the wall which is protected by a steel door; behind this is a heavy security rubber door which only opens inside to admit security materials, and which shuts immediately afterwards. In cases of emergency, however, a door linking the security room with the loading area may be used, but it can only be opened from inside and the key is kept in the strong room. A small gravity-type conveyor belt transfers materials rapidly from the loading area to the security room.

Both the strong room and the script room can be entered directly only through the security room. However, materials can be passed back and forth between the data processing department and the script room through the common room between them. Double doors separate the security room from the other two rooms; the inner door is steel and the outer one is wood.

1. Keys. The main door to the security room, which is made of steel, is locked with a combination lock, the combination for which is shared between the assistant directors of office operations and test administration. The statement of the combination formula is also placed in a sealed packet and placed in the director's safe. A set of keys of all other locks leading to the security rooms is held by the security officer, with a duplicate set in the director's office safe.

The security officer must keep his complete set of keys in a safe, and whenever he leaves the office, he should give his security keys to an assistant director and record the transfer on a key register (Figure 9.7).

The stores officer has a key to the script room door, and the data processing manager has a key to the door in the data processing department which leads to the common room. While scripts, answer sheets, certificates,

Figure 9.3 Stores register.

Description:	Consignor receipt no.	Date	STO inventory control no.	Consignment accepted by	Remarks

Figure 9.4 STO inventory control

Number _____

Description:	Quantity	Date	Location	Officer	Remarks

Figure 9.5 Shipping register

Description:	STO inventory control no.	Date	Consignment released by	Consignment accepted by	Remarks

Figure 9.6 Access register

Date	Time in	Time out	Nature of transaction	Signature in	out

Figure 9.7 Key register

Key number	Usual holder	Transferred to	Date/time	Signature

and other confidential materials can be passed back and forth via this common room, personnel are not allowed to use it as an accessway.

Under no circumstances whatsoever are keys to the security rooms taken out of the testing office.

2. Access. Access to the security areas is strictly limited, and either the assistant director for office operations, the security officer, or the stores officer must be present all the time the door is unlocked. A security violation will be charged against any of these three individuals who leaves his post without making certain that one of the others is present, or that the door is locked.

The officer taking charge of the operations should take into the security area only the minimum number of staff members to assist him. As packets of paper are unpacked here they should be checked for evidences of tampering, counted, and passed into the security room where the changes in number of packages and other data are entered on the STO inventory control card, Figure 9.4. The control card shows the number of separate packages in storage at any one time. The shipment may have consisted of one large crate in which there were 60 individually wrapped packages. The card would therefore show one crate on the first line, which when unpacked shows on the second line as 60 packages. The officer's responsibility is to determine by a physical check that the balances he records on the cards at the end of each operation are correct. (An operation is ended each time the security area is vacated and locked up. Whenevery this happens, the contents of the security area consist only of unopened boxes and what is locked up in the security room.)

Except for the three officers mentioned above, and any other officers equal to or above the grade of assistant director, everybody who has been allowed into the security room must be searched upon coming out.

The security officer is also responsible for seeing that everyone signs the access register, (Figure 9.6), in the security room. This room is off limits to all officers except those on business, and everyone entering it must sign this register, *regardless of his title or seniority.*

Anyone who accompanies either the security or stores officer to the security room must sign the register, stating his time of entry and departure, and the nature of his transaction. Staff members working within the room should sign even when they leave for short routine absences, such as coffee

and lunch breaks, or trips to the lavatory.

The assistant director for office operations must inspect the security area from time to time when operations are in progress to satisfy himself that the registers and inventory control cards are being properly maintained and that the balances shown thereon are actually present. That is, any materials included on the stores register but not included on the shipping register (Figure 9.5) must be present in the security area, and their exact location ascertainable by reference to an inventory control card.

The assistant director for office operations must inspect the security area at least once a week, and make a formal, detailed report to the director, and both the assistant director for test administration and the assistant director for data processing should visit the security area occasionally when operations are in progress to check that regulations are being observed and instructions carried out.

B. Safes. The safes are assigned to assistant directors by the director or his designate and are kept in the regular offices. General regulations pertaining to safes are given below:

1. Any assistant director who is responsible for handling classified material will be allocated a safe for his office.

2. The transfer of the safe will be registered by the security officer who will record that the assistant director is in possession of one of the two keys to the safe. A copy of the record will be kept in the safe of the assistant director for office operations.

3. Any assistant director who has been allocated a safe will be responsible for having it locked at all times. If any officer is found by the security officer, or any other assistant director, to have left the safe open while not actually depositing or extracting materials from it, he shall be reported to the director who will take disciplinary action against him for breach of security regulations. The misdemeanor is to be duly noted in the individual's personnel file.

4. The key of any safe must always be in the secure possession of the officer responsible for that safe. Under no circumstances may such a key be deposited in a locked drawer or in any place less secure than a steel safe.

5. If an officer to whom a safe has been allocated goes on leave or is absent from the office for a lengthy period of time and another officer is appointed to perform his duties during his absence, the key of his safe

should be handed over to the officer replacing him. The assistant director of office operations will then amend the register of officers in possession of safes in accordance with the new arrangement. Officers holding safes must maintain an up-to-date inventory of the contents, and copies of this inventory, together with the key, must be handed over to any replacement in the presence of either the security officer or the assistant director for office operations after both the outgoing and incoming officers have jointly checked the contents and signed the inventory.

6. If any officer is going to be absent for such a short time that no other officer is appointed to take his place, he should give the key to his safe and the inventory of its contents to the assistant director for office operations and notify the security officer of his action.

7. The duplicate key of each safe registered as having been allocated to an officer will be kept in the safe of the assistant director for office operations, and will be listed on the inventory of the contents of *his* safe. The inventory will be amended whenever the registry of officers entitled to be in possession of safes is amended by the director, or a key is temporarily given to him (as in number six above).

8. The duplicate key to the assistant director for office operations' safe shall be kept in the director's safe.

9. Assistant directors should lock the door of their offices after office hours. If they intend to be away from the office for more than a few days, keys to their offices should be left with officers of equal or higher rank.

II. Receipt of Materials.

Every single crate, parcel, package, envelope, or box--every single item which enters the security areas must be acknowledged by a receipt, kept under a running inventory count, and dispatched with a receipt.

Either the security officer or stores officer must be present when a shipment is received. He must satisfy himself that the container is intact; if it is, he completes an entry in the stores register, signs the consignor's receipt if necessary, and stamps an inventory control number on every container in the shipment. Since the control number is shown on the stores register, the consignment can be located within the security areas when necessary. In the space for comments, he should note any circumstances which seem suspicious; e.g., any unaccountable delays, the fact that the outer

wrapper is intact but the inner one is damaged, that some papers are out of order, that one paper is exceptionally soiled, that unusual tool marks appear on the corners.

If the container is not intact, the contents must be carefully removed, inspected, and cleared to the appropriate security area immediately. Each separate package must first be stamped with the inventory control number before transfer to storage, and the number of sub-parcels noted on the stores register.

In addition, an inventory control card (Figure 9.4) should be filled out for every shipment. The same inventory control number stamped on the stores register (Figure 9.3) must be stamped at the top of the inventory control card, and the description of materials entered. The card will be used to keep an account of the number of packages, the dates that the packages were broken down to smaller units or repacked as larger units, their physical locations within the security areas, and the signature of the officer who made the particular change, or who moved the materials.

If security material arrives by airfreight or rail, the security officer is responsible for its immediate collection, and makes certain that a suitable officer with appropriate experience and seniority collects the consignment. The officer collecting the consignment is responsible for its arriving complete and intact and *does not allow it to be opened by Customs or other officials on any pretext.*

If the parcel is too bulky to be kept in a safe, the collecting officer arranges with the security officer for the security room to be opened immediately upon his arrival, whatever the time. The security officer checks the consignment against the waybill, and when he is satisfied that the quantity is correct and the sealing is intact, supervises the off-loading of the consignment and makes the appropriate entry on the stores register.

All consignments must be handed over by the collecting officer to the security officer immediately after arrival at the office. In no circumstances may any consignment or part thereof be left even temporarily in an office, car, lorry, house, or other non-security area.

If security material arrives by registered mail, it is handed personally to the assistant director for whose attention it is marked, or in his absence, the security officer places it in the safe until it can be delivered to the appropriate officer.

III. Storage.

Depending on the flow and quantity of work awaiting completion, incoming goods will be transferred intact to the strong room until action can be taken, or moved to the security room for unpacking and repacking. Whatever action is taken must be noted on the inventory control card.

IV. Packing.

Almost all materials which come to the security area are repacked and shipped to other places; the remainder are shredded and burned (Details of the burning operation are given later in this section.) Packing takes place primarily in the security room under the direction of the assistant director for office operations.

All packing is carried out with securitape sealing. Securitape is a special, colored cellotape with the STO's initials reproduced every half-inch along its length. When it is removed from a package, part of the colored coating, along with the initials, flakes off preventing any re-use. The blue securitape might be used for all live question papers, manuscripts, answer keys, proofs, and any other confidential documents relevant to live question papers. Other security materials of a confidential nature could be sealed with the black securitape.

All tins containing rolls of securitape should be kept in the strong room until they are opened. Once opened, the individual rolls must be kept in a locked safe, the key to which is held by an officer not below the rank of assistant director. Under no circumstances should securitape be left lying around in drawers or unlocked cabinets. The only officer below the rank of assistant director allowed to use the securitape at all is the security officer. After use, envelopes and packets sealed with securitape should be cut into very small pieces or shredded.

Bulky parcels of outgoing question papers must be packed in cardboard containers, wrapped in stout brown paper, and sealed with securitape. The outside of each parcel should be marked "CONFIDENTIAL" and bear the name of the person to whom the parcel is being sent, as well as the name of the center or hall where the papers are to be used. The parcels should then be enclosed in additional sheets of brown paper, crossbanded with a metal strip, and sealed with a metal seal. The outer paper should bear the designation or name of the person to whom it is being sent, but no indication that the contents are of a confidential nature.

If the number of packets of question papers to be dispatched is small, they may be placed inside an envelope together with the consignment note. This envelope, marked "CONFIDENTIAL," should be sealed with securitape, addressed by name to the person for whom the question papers are intended, and enclosed in an outer envelope bearing the name or designation of the person to whom it is being sent, but no indication that the contents are of a confidential nature.

All envelopes containing security materials (such as inner wrappers of material being dispatched and wrappers of material being held in safes) must be sealed with securitape. The method of sealing is as follows:

a. One or more strips are positioned across the flap of the envelope and aligned exactly across the top edge of the flap, overlapping the edges of the envelope at least one inch.

b. One strip is placed across the bottom joint of the envelope, similarly positioned and overlapping the edges.

c. One strip of securitape is located across each vertical joint of the envelope, overlapping the edges by at least one inch.

d. In the case of envelopes which are to go in the post and which are not of stout material (although the use of these is to be avoided), one strip should be folded over each edge of the envelope.

Before he seals with the securitape as outlined above, the officer must sign his name or initials at the seams of the envelope.

Examination booklets, answer sheets, and other testing materials are ordinarily packaged either by the printer or at the STO in units of 5, 10, 25, 50, and 100 approximately in the ratio of 16:8:4:2:1; (i.e., there should be twice as many packages of 5 as of 10, twice as many 10 as of 25, etc.)

The quantity of papers supplied to a center should be equal to the number of candidates entered, plus enough to round up to the next five; for example, a center with 42 candidates would get 45 papers, one with 45 would get 50. Unless there is a shortage of question papers, envelopes or cartons of papers packed and sealed by the printers should not be repacked in the office. If it is necessary to repack, the quantity of papers supplied to a center need not be rounded up to the next five. *Under no circumstances, however, should a whole packet or carton of question papers be forwarded to a center in excess of the number of candidates entered.*

V. Shipping.

All shipments or other removal of confidential material from the security room *must* be noted on the shipping register (Figure 9.5). At the time of the shipment, the STO inventory control card should be removed from the active file, dated, and placed in the inactive file with the notation that the new location is now "external."

Before the examination, supervisors will receive question papers in one of two ways: (1) the question papers may be handed over personally by a STO local representative, or special education officers and supervisors may be asked to collect them at stated collection centers at prearranged times; or (2) the question papers may be airfreighted, trucked, railed, or otherwise sent to the special education officers.

Details of what is enclosed in each sealed packet of question papers are marked on the inside envelope. Under no circumstances (but one) may the envelopes be opened until immediately prior to the time scheduled for the examination to begin. The only exception to this rule occurs when certain practical or oral test papers must be given to candidates well before the actual examinations; envelopes containing such special papers may also be opened, repacked, and sealed, also under the strictest supervision. Special instructions to this effect will be indicated clearly on the question paper envelopes.

When he receives the test papers, the special education officer checks the stated contents of the sealed packets endorsed on the outside of the envelope against the invoice or waybill. He also checks the condition of the sealed packets, satisfies himself that they are intact, and signs the invoice or waybill. Any shortages or badly sealed packets must be mentioned to the delivering officer and recorded on the invoice or waybill. Badly sealed packets should be handed back to the delivering officer, who notes this on the documents or has them properly resealed. Where the question papers arrive by airfreight, the security officer at the STO should be notified by telegram or telephone.

Where there are two copies of the invoice, a copy should be given to the delivering officer. Where there is only one copy, the special education officer signs and hands to the delivering officer the following statement: "I have received intact all the question papers listed on the invoice with the following exceptions . . . ." It is necessary that the SEO keep a copy

of the receipted invoice in safe custody so that it can be made available at any time for inspection by the supervisor and visiting Ministry of Education or STO staff.

If question papers are to be delivered personally, a member of the testing office staff, preferably not below the rank of examinations officer, or an assistant education officer in the ministry, or a lecturer at the university, should be chosen to do it.

Armed guards must be employed whenever question papers undergo a journey--other than by post--longer than one day, whether accompanied by STO staff, or by special education officers and supervisors.

In any case where materials are distributed with the expectation that they will be returned, the shipment should include a strong envelope of appropriate size or brown wrapping paper sufficient to serve as the inner wrapper; this inner wrapper is marked "CONFIDENTIAL" and addressed to the assistant director concerned. Another strong envelope, longer than the other, or additional brown wrapping paper, should be included to be used as the outer cover for the security materials. This outer cover is preaddressed to:

The Director
State Testing Office
P.O. Box 1234
Regina, Atlantis

and should bear no mark to distinguish it from other mail.

Security material to be posted must be registered, and the registration particulars copied on the file, float, and advice copies of the covering letter by the officer dispatching the packet.

The sealed, addressed packet as described above must be handed personally by the examinations officer to the dispatch secretary who makes the necessary entries in the dispatch register. The packet must not be placed in an "IN" or "OUT" tray. The dispatch secretary personally applies the correct postage and hands the packet, plus the registraion book, to the assistant director for office operations who keeps it in his safe until such time as he or the security officer is personally able to post it at the post office.

On occasions security material may have to be sent by airfreight. In such cases, the assistant director concerned makes sure that the parcel is well packed, sealed, and addressed before it is dispatched, and that an advice note is sent to the addressee. The assistant director personally hands the parcel to the dispatch secretary to send on the first available

plane. The dispatch secretary subsequently sees to it that the freight waybill and associated papers are given to the assistant director, who communicates these by cable to the addressee.

At least once each year the assistant director for office operations consults with the director and decides which security materials are ready for the ultimate journey: to oblivion. A letter drafted by him and signed by the director will be accepted as authorization to take the material out of the security area, through a shredding machine, and from there to the incinerator. The security officer must make certain that every scrap is burned and that the ashes, which can retain an impression of the printing, are thoroughly crushed and scattered. It is not sufficient to tear up unwanted security material, or merely to set fire to it. Paper burns slowly and incompletely if left to itself, so the security officer has to supervise the burning from the first flame to the last scattered ash.

VI. Test Centers.

Before and after the examination, the special education officers, STO local representatives, or supervisors should keep the question papers in steel safes or any of the following strong rooms:

1. the vaults of a sub-treasury.
2. the vaults in a bank.
3. the strong room at a police station.
4. any other room approved by the STO as sufficiently safe, having steel bars in the windows and strong locks on the doors.

Whenever possible, the question papers should not be issued to the supervisor until the day of the examination. But if for some reason the supervisor is given the question papers before the day of the examination, he must keep them in complete security in a steel safe. Supervisors should be instructed that as soon as they have received the papers from the special education officers or STO local representatives, they are responsible and accountable for their security (see page 7-10).

The sealed packets containing the question papers must be opened only in the presence of the candidates just before the examination is scheduled to commence. Because the STO requests the return of the original envelopes along with the answer sheets in order to check for tampering, among other things, the envelopes should be cut neatly along one side, not just torn to pieces. In cases of any question about security, the name of the

officer who packed the envelope can be ascertained from the envelope itself, which he signed several times. He provides a reasonable starting point for an investigation. This envelope should also be signed by all taking part in the supervision or invigilation of the paper concerned, certifying that the envelope was intact when received and that it was opened in the presence of the candidates just before the examination began. A specimen envelope sealed in the required way should be sent in advance to each center with the instructions to supervisors and invigilators. See Figure 7.1 for a certificate of examination conduct.

Outside the offices, security can be improved if staff of the Ministry of Education, special education officers, and STO local representatives inspect centers during the examinations.

[At examination centers candidates should be searched, in cases of reasonable suspicion, before entering the examination room so as to prevent the smuggling of prepared answers into the room. Female candidates must only be searched by a female supervisor or invigilator. In cases of reasonable suspicion, candidates should also be searched upon leaving the examination room, from which they may try to remove answer booklets, used or unused, and notes made during the examination.]

Supervisors are required to complete a seating plan for each examination session during that session and to forward this with the respective answer papers to the State Testing Office.

After the test, the supervisors must wrap the candidates' completed answer sheets carefully in brown paper and include in the parcel a certificate of conduct which they have signed, secure the parcel with string, and seal it with wax; there should be no indication on the parcel that the contents are of a confidential nature. The only way to make certain that the supervisor will do the wrapping properly is to send preaddressed and prepared materials to him along with the tests. He should be instructed to pack examination materials as stoutly as possible, since parcels are subject to rough treatment by postal authorities and can be badly damaged by the time they arrive at the STO. A copy of the consignment note should be sent by the quickest means to the Director.

The parcels containing answer sheets should be dispatched immediately by airfreight when this is possible; when it isn't, the next best route is by registered post. Where it is not possible to dispatch the answer sheets

immediately after the examination, the supervisor must keep them completely secure in a safe or in a locked steel cabinet.

The special education officers or STO local representatives may ask the supervisors from some centers to deliver the answer sheets directly to them for onward transmission to the STO. In this case, the same precautions should be observed. (No matter what the extent of precautions, the test booklets are no longer secure after the examination is completed. The author has never seen a large-scale examination where every booklet that was sent out was returned, but see page 7-6 for the considerations involved.)

VII. Data Processing Department.

The data processing department is adjacent to the security areas and during the scoring/processing/reporting phase is also considered to be a security area. During this time the data processing manager uses his key to obtain access to the common room where the security officer or stores officer transfers materials to him.

Access to the data processing department is strictly prohibited except to the data processing manager and his regular staff, the computer engineers, and the STO maintenance staff (only when required to clean the room or to inspect any part of the plant). When the latter personnel are on the premises, they must be constantly under the watchful eye of the data processing manager. Permission for other STO staff or for outside visitors to enter the department must be obtained from the assistant director for data processing.

All punched cards, magnetic tapes, magnetic disks, data printouts, and any answer sheets still being processed must be stored in the common room when they are not being used.

The doors to the data processing department are locked whenever no one needs access to the department, and the key is held by the data processing manager or assistant director for data processing.

Steel cabinets, which are used for storing security materials in the data processing department, including the punch section and the data processing department itself, must be locked at all times outside the office hours except when members of staff are working overtime under supervision.

Answer sheets are transferred in bulk from the script room to the common room where they remain until scoring begins. Where necessary, the data processing manager allocates the answer sheets or mark listings to

punch operators and records the name of each operator who punches or verifies each batch of data.

After being scored, the answer sheets are returned to the script room via the common room, and the cards are kept in the common room files until they are required. The mark listings are also kept in the common room until the examination processing is completed.

VIII. Typing, Copying, Printing.

The typing of question papers and other security materials, whether in draft or final form, is done only by the confidential secretary. The assistant director responsible for the materials ensures that all possible precautions are taken to preserve the security of the material during this operation; in particular:

a. that the typist is given a room in which she can lock herself and where she cannot be disturbed.

b. that all manuscripts, spoiled copies, carbons, carbon ribbons, etc. are handed over to him when the typing is completed.

c. that the typist complete a certificate, stating that the documents have been prepared under security conditions and that all materials have been handed over to the assistant director.

Photocopying of draft or final question papers is sometimes necessary in cases or emergency, or for pretesting. Whenever possible, the assistant director *personally* does the photocopying and duplicating of draft question papers. If this is not possible, he closely supervises the work of a junior officer, never leaving him alone during the process and always insisting that the principles laid down in the preceding paragraphs be followed. After the copying process is completed, the assistant director signs a security statement and also obtains the endorsement of the junior officer.

The copying and printing must be carried out behind locked doors, and where practicable, in the security room. The room used for the process, or at least an appropriately large area of it, must first be cleared completely of all other papers; drawers of duplicating machines must be emptied.

The backing sheets of stencils and spoiled copies are collected as they are produced and stored in a large envelope. To such material are added the stencils themselves after copies are made. A thorough check of the duplicating machine should be made to be certain that no paper or copies remain in it, and, so that they will not carry an impression of the stencil, the rollers are

thoroughly cleaned. If a carbon ribbon was used in typing, it is removed and placed in the envelope as well.

Collating and stapling of confidential materials should also be done in the security room under the standard security regulations. Partially or badly printed copies are sealed in envelopes or boxes and locked up in the strong room until the end of the examination processing.

IX. Confidential Files.

Regulations governing the use and transmittal of confidential files are given below:

1. All confidential files are to be placed in red covers and clearly stamped "SECURITY" on the cover.

2. Only the assistant director and the director may have access to confidential files. An assistant director may authorize an officer not below the grade of examinations officer to have access to any file to which he himself has access, provided he is satisfied that this is essential for the performance of that officer's duties.

3. Certain confidential files dealing with examination policy or other security matters concerned with a particular examination (appointment of examiners and minor irregularities, for example) may be kept by the secretary/typist of the appropriate assistant director. With respect to these particular files, the secretary/typist concerned is required to follow all security regulations relating to the safekeeping of security material. The assistant directors must give the security officer a complete and comprehensive list of such files kept by their secretary/typists. This list will be examined by the assistant director for office operations who will decide whether or not any particular security file should be kept by the secretary/typist concerned. It is, of course, evident that no officer has access to his own confidential personnel file or to any confidential file of any officer of equal or senior grade.

4. No officer may leave his office if a confidential file or other confidential material remains on his desk or elsewhere in the office other than in a locked safe.

5. All security files and confidential material must be locked in the safe or returned to the assistant director for office operations if the officer concerned leaves his office no matter for how short a time. It is not sufficient to lock the door of the office.

6. The security officer and assistant director for office operations should undertake frequent inspections of all offices both during and after office hours to check that no confidential material has been left accessible in an office in the absence of the occupant.

7. Any officer who entrusts a file to a secretary/typist in no way absolves himself of the responsibility for its security. If he wants to pass the file to another senior officer who is authorized to have access to it, he must first have it returned to him and then personally pass it on.

8. It should be evident that under no circumstances can anyone, except a person authorized to handle the particular security file concerned, carry that file from one office to another for anyone else.

X. Security Breaches.

A little experience will show that no matter how carefully the security regulations are drawn up, a breach will eventually occur. Of course, the most effective deterrent is close supervision since the fear of being caught is a better preventative than the severity of punishment. If, however, the general attitude of the senior staff is that all headmasters, supervisors, candidates, and lorry drivers are gentlemen and should therefore never be suspected of anything as gross as thievery, no such thing as adequate security can exist. Once when the author tried to cajole his colleagues into some spot inspections at various testing halls, he was rebuffed with the "gentlemen" argument. They were almost correct; only one case of leakage occurred that year. One method worth trying is to choose locations at random, and assign these at random to *all* senior officers with instructions to complete a security and administration report on the centers.

Anyone reading between the lines of the regulations will know the types of security breaches which have commonly occurred in the past. A few specific instances should help to prove that the unscrupulous, like the poor, are always with us.

A special education officer was contacted by telephone by a man claiming to be the assistant director of the State Testing Office. The man said that about one-half of the test papers in one subject had serious typographical errors, and advised the officer that he had official permission to remove the test papers from his safe, open the sealed packet, and read paper 317/A to the "assistant director" who would then give the correct wording if the paper was defective.

One year all papers were sent in special window envelopes in which the name of the test was visible through a cellophane-covered opening. The edges of the envelope were carefully sealed with securitape. A cunning supervisor merely removed the cellophane, rolled one paper around a pencil inserted through the opening, and removed the paper. After copying it, he inserted it back into the envelope, glued the cellophane into place, and kept officials wondering for a very long time how he had accomplished his feat.

The State Testing Office often unintentionally asks for trouble. Once, when working with a small budget, they asked certain schools to collect particular types of rock specimens in anticipation of a national examination, and to distribute them to other schools immediately prior to the examination date. Naturally the select schools primed their students on everything related to the types of rocks they had collected, and carefully refrained from passing on any critical information to the other schools.

In another country, the headmasters were assigned as invigilators to schools other than their own because the STO suspected that otherwise collusion might take place. After taking such painstaking measures to protect the examination from the headmasters, however, it turned out to be a supervisor who opened the package in advance, gave a copy to the headmaster who would work as his invigilator the next day, and took a second copy for his son.

One enterprising young man simply asked his brother in the data processing department to add his name to the general lists of candidates who passed the test. Luckily for the STO, the count didn't agree and the brothers shared a year in jail.

Another time, the STO's gentlemanly driver picked up 20 crates of examination papers at the docks and delivered 19½ to the office. Investigation disclosed that he was a candidate for the examination and just couldn't resist the temptation. Now the office has every employee sign an affadavit that he will inform them of any examination he is preparing to take.

A senior officer at the STO put a dozen packages of examination papers in his car in front of his home for storage overnight before leaving in the morning for a distant center. When he arrived at his destination, he had only eleven. Unbeknownst to him, his nephew who was staying with the family, had the twelfth.

Some plots don't involve leaks at all. A group of candidates, worried

with good reason about their ability to perform on an examination, decided that if they had too much trouble with one paper, they would cry out "security leak." So when they inevitably realized that they had no hope of passing the paper, they proceeded to carry out their scheme, complained vociferously that the examination had leaked over a week before, and demanded that another examination be set. Fortunately, their invidious plot was uncovered in time.

Anyone who has ever worked with tests can add a dozen or more cases to those presented here, and knows that to draw up a reasonable list of security regulations and procedures, all an STO officer has to do is to imagine every possible devious method of obtaining a copy of the papers, assume someone will try it, and take steps to make it impossible. Chances are that someone will try something devious, and will get away with it, unless he has been anticipated, or unless the testing office is having a most uncharacteristically lucky season.

COSTS

Typical costs for each stage of the examination cycle are difficult to obtain since very few governmental agencies use cost accounting methods. Civil servants who take part in the examinations are reluctant, even if they have accurate information, to disclose the time they spend on duties ranging from writing directions to meeting with irate parents.

However, the officers in charge of testing in Malawi, Botswana, Lesotho and Swaziland were interviewed in 1970 concerning the various costs for their large-scale examinations, and part of the data will be presented here. These four countries are among the least wealthy in Africa, and since no estimates of officer's time could be obtained, the expenses can be regarded as a lower limit. Each stage of the examination cycle will be briefly considered, and the range of identifiable costs among the four countries given in U.S. cents per candidate.

Publicity. No country offered a brochure or advertised the examination in the press. All gave information concerning entries and dates either through a ministry circular or letters to the District Education Officers; two countries provided model papers to the candidates. The only cost estimate was 0.7¢ per candidate.

Registration. Three of the ministries tested candidates at the end of standard seven, and other at the end of standard eight. The number of candidates ranged from about 5,000 to 27,000, and all candidates were registered through their headmasters. No information was available on typical costs.

Item Preparation. The questions were set by education officers, inspectors, headmasters, and teachers with competence in the subject matter area. In the country where formal training of examiners was conducted, it cost 3.5¢ per candidate, but in the others where an older examiner trained a younger one, the cost varied between 1.7 and 8.7¢ per candidate.

The examinations were a combination of essay, short answer, and multiple choice questions in four areas: English, arithmetic, general knowledge, and either the local language or science/housecraft. Candidates were typically allowed slightly over an hour on each paper.

Pretesting. Only in one country were the questions pretested. The items were sometimes added at the end of the regular examination and the enlarged papers administered to selected schools, and sometimes tested

separately a few days after the regular examination was finished. No cost estimates were given.

Editing. The final editing of the items and preparing them for reproduction was done either by a committee, or by officials from the Ministry of Education. The country which budgeted for this step averaged 0.7¢ per candidate.

Reproduction. Costs per candidate ranged from 10.3 to 20.5¢, and all were based on offset reproduction of the tests and answer sheets. Two of the countries utilized the services of the government printers, one the services of the local correspondence college, and the remaining one the services of a printer outside the country (with the highest cost as a result.)

Distribution/Administration. The distribution of the examination papers was by a combination of post and motor vehicle, never by rail or air. Costs of distribution and return of materials were between 1.8 and 10.7¢ per candidate, with the differences depending on the form of transport and distances involved.

The variety of persons conducting the examination was considerable: headmasters, examination officers, educational officers, teachers, local politicians, and parents. Typically the examination took three mornings to administer, and cost between 0.9 and 12.0¢ per candidate, including administrators' incidental travel costs. No information was available concerning the expenses of printing the directions.

Scoring/Processing/Reporting. Scoring costs were considerable except in the one country which used the computer for this step. To underscore some of the earlier remarks about preferences for using a machine, consider the following reports: 50 teachers mark for one week with an unknown error rate; 240 teachers mark for two weeks with an unknown error rate; 12 inspectors and 48 teachers mark for three weeks with an estimated error rate of ten per cent. (The estimate of errors was refreshingly honest.) Costs were 88.1, 7.5 and 17.5¢ per candidate respectively. In the years before transferring to the computer, the government of one country had employed 300 teachers for two weeks at a cost of 20.3¢ per candidate; after the move they employed 30 teachers for the essays at a cost of 8.6¢ for both teachers' and machine scorings combined. One might ask what cost figure is a fair representation of the expense of teachers' time that

they might otherwise devote to study and course preparation.

Processing of the marks, which included transformations, totaling for each candidate, assigning grades, sorting to order of merit or whatever the ministry required, and printing of marks and certificates, cost between 7.1 and 28.0¢ per candidate. All four of the countries utilized the computer for this step.

Printing of the certificates, at a cost between 1.2 and 3.0¢ per candidate, was low, compared to West Africa for instance, largely because the printing was done within each of the four countries.

Total Costs. The total cost per candidate varied between 49.0¢ and $1.69, with no consistent relationship between the number of candidates and the per candidate cost; i.e., the cost was not necessarily less per candidate where more were tested at one time. As mentioned earlier, the estimates do not include manpower costs for the many civil serviants who assisted throughout the examination process, nor for the teachers who generally were not paid for their marking. In comparison, marking of scripts alone in the U.K. was estimated to cost 47¢ each (School Council Examination Bulletin # 12, 1966, p. 26).

The percentage of candidates which was accepted into secondary school was 50 per cent in the country with the highest selection ratio and 12 per cent in the country with the lowest. When only a small percentage is chosen, it is imperative that the waste of human resources be kept to a minimum. To the extent that testing reduces such loss in secondary school, it justifies the examination costs involved.

One situation mentioned in the preface might be expanded. When the Polytechnic College of the University of Malawi was ready to begin instruction, applicants for the 60 places were received from 1,350 applicants. The number was reduced by a three-stage process: (1) elimination of the 1,150 candidates who received less than straight A ratings on character, scholastic ability, and attendance from the primary schools they attended; (2) selection of the best 120 of the remaining 200 by a central review panel; and (3) identification of the 60 out of the remaining 120 by use of English and arithmetic tests followed by an interview.

In spite of the care exercised, at the end of the first term five had been dismissed, six had been placed on probation, and nine were given severe warnings--fully a third of the entrants were unsatisfactory.

The high cost of failure in countries where educational facilities are seriously circumscribed more than offsets the expense of initiating or continuing large-scale examination programs. Carefully conducted, valid examinations should be regarded as an economic necessity.